Technische Universität Darmstadt

Fachbereich Rechts- und Wirtschaftswissenschaften

Meta Analysis of Crime and Deterrence

A Comprehensive Review of the Literature

Vom Fachbereich genehmigte Dissertation
zur Erlangung des akademischen Grades
Doctor rerum politicarum (Dr. rer. pol.)
vorgelegt von

Dipl.-Math. Thomas Rupp (geb. in Montevideo, Uruguay)

Referenten:
Prof. Dr. Horst Entorf (Erstreferent und Betreuer)
Prof. Dr. Dr. h.c. Bert Rürup (Koreferent)

Tag der Einreichung:
31. März 2008

Tag der mündlichen Prüfung
1. Juli 2008

Darmstadt, 17.7.2008, D17

In memory of my mother.

Thomas Rupp

Meta Analysis of Crime and Deterrence

A Comprehensive Review of the Literature

Contents

List of Tables

List of Figures

List of Abbreviations

AIC	Akaike Information Criterion
ANCOVA	Analysis Of Covariance
ANOVA	Analysis Of Variance
ARIMA	Autoregressive Integrated Moving Average
BAC	Blood Alcohol Content (or Concentration)
BIC	Bayesian Information Criterion
BMA	Bayesian Model Averaging
BMS	Bayesian Model Selection
CDC	Centers of Disease Control and Prevention
CDF	Cumulative Distribution Function
CSS	Cascading Style Sheets
df	degrees of freedom
DFG	Deutsche Forschungsgesellschaft
DUI	Driving Under the Influence
DUII	Driving Under the Influence of Intoxicants
DWI	Driving While Intoxicated (or Impaired)
EBA	Extreme Bounds Analysis
ECM	Error Correcting Model
FBI	Federal Bureau of Investigation
GDP	Gross Domestic Product
GLS	Generalized Least Squares
GMM	General Method of Moments
HTML	Hypertext Markup Language
IRS	Internal Revenue Service
LPS	Log Predictive Score
MADD	Mothers Against Drunk Driving
MANOVA	Multivariate Analysis Of Variance
ML	Maximum Likelihood
MMA	Mallows Model Average (Estimator)
NCHS	National Center for Health Statistics
NCVS	National Crime Victimization Survey
NHL	National Hockey League

NLSY	National Longitudinal Survey of Youth
NRA	National Rifle Association
NYS	National Youth Survey
OLS	Ordinary Least Squares
PHP	PHP: Hypertext Preprocessor
PKS	Polizeiliche Kriminalstatistik
RMSE	Root Mean Squared Error
RMSPE	Root Mean Squared Proportional Error
RSDA	Rough Set Data Analysis
sd	standard deviation
SUR	Seemingly Unrelated Regressions
UCR	Uniform Crime Report
VAR	Vector Autoregression
WHO	World Health Organization

Index I crimes	murder and non-negligent manslaughter, forcible rape, robbery, aggravated assault, burglary, larceny-theft, motor vehicle theft, and arson
Index II crimes	all crimes except Index I crimes and minor traffic violations (amongst others: simple assault, narcotics, vandalism, vice, fraud, major traffic violations)

Preface

The average Ph.D. thesis is nothing but a transference of bones from one graveyard to another.

Frank J. Dobie, A Texan in England, 1945

We have a habit in writing articles published in scientific journals to make the work as finished as possible, to cover up all the tracks, to not worry about the blind alleys or describe how you had the wrong idea at first, and so on. So there isn't any place to publish, in a dignified manner, what you actually did in order to get to do the work.

Richard Feynman, Nobel Lecture, 1965

After several years of planning, reading, calculating - fixing mistakes, recalculating, updating data and redoing all calculations all over again a few times - and writing, this comprehensive meta analysis is finally ready to be published. Although it would be misleading to call it finished, I am sure that it covers all necessary aspects to stand on its own. It would have been very easy for me to extend this work almost to infinity by including more studies, performing more calculations, applying more techniques or spending more time and space on many aspects. Eventually, even for a literature focussed thesis with a large statistical coverage it has become very long. I conceptualized this work in such a way that most of it may be read and understood by almost any interested reader; only a few sections require some advanced statistical knowledge to be fully comprehended. Some readers who are already very familiar with some concepts may want to skip the corresponding sections.

This work emerged from a cross-disciplinary project between the economics department of the Technical University of Darmstadt and the criminological department of the University of Heidelberg. The project started in 2003 and was financed by the Deutsche Forschungsgesellschaft (DFG) until 2007. I picked up many ideas, elements and conceptions and introduced them into this thesis. Although I use the first person plural throughout the whole thesis, this has solely stylistic reasons. Whenever I refer to multiple persons, I do this explicitly.

Some quotes introducing new sections are missing a primary source. Although the given author was verified as well as possible, I had to rely on secondary and internet sources. This work includes a very large appendix which contains some typical elements which are not necessary for understanding the thesis. However, most of the appendix is made up of the descriptive coverage of all included variables and studies. These are not essential but ultimately belong to this meta analysis.

Many people were - directly and indirectly - involved in the making of this thesis. The team in Heidelberg consisted of Dieter Dölling, Dieter Hermann, Andreas Woll and Armando Häring. The team in Darmstadt was made up by Horst Entorf and myself. Needless to say that my acknowledgment of their support and work are placed here first. I am especially grateful for the advises, resources, patience and academic freedom given to me by my PhD supervisor Horst Entorf. I have also to emphasize the technical support by Ileana Petroniu who was an indispensable help in processing and cleaning the literature data base, acquiring new studies and other administrative tasks; as well as Philip Savage for many advises on the English language. Furthermore, I have to thank Hannes Spengler, Oliver Schmid, Emanuela Trifan and Jochen Möbert for suggestions and uncovering mistakes. I also incorporated many suggestions made during several criminological and economical conferences.

1 Introduction

Without libraries what have we? We have no past and no future.
Ray Bradbury

Meeting Dr. Wollaston one morning in the shop of a bookseller, I proposed this question: If two volumes of hydrogen and one of oxygen are mixed together in a vessel, and if by mechanical pressure they can be so condensed as to become of the same specific gravity as water, will the gases under these circumstances unite and form water? "What do you think they will do?" said Dr. W. I replied, that I should rather expect they would unite. "I see no reason to suppose it," said he. I then inquired whether he thought the experiment worth making. He answered, that he did not, for that he should think it would certainly not succeed. A few days after, I proposed the same question to Sir Humphry Davy. He at once said, "they will become water, of course;" and on my inquiring whether he thought the experiment worth making, he observed that it was a good experiment, but one which it was hardly necessary to make, as it must succeed.
Charles Babbage, Reflections on the Decline of Science in England, and on Some of Its Causes, 1830

Crime matters to society. As reported elsewhere[1] - and according to common sense - crime causes huge economical and psychological damage to individuals, the economy and to society itself. Therefore, it is natural that a society - through its legislative, executive and judiciary - tries to control crime. For thousands of years the idea that the fear of arrest and subsequent punishment will deter people from committing crimes, has been a major tool in this concept. While punishment has several motivations such as revenge, retribution, normative guidance, correction and deterrence, the latter was theorized late in the 18th and 19th century[2] by Marchese Beccaria (1819); Bentham (1830) and Chadwick (1829). Deterrence is recognized as a method for preventing potential delinquents[3] from committing crimes by the threat of punishment. As far as deterrence is concerned, punishment is not meant to be anything like a "fair compensation" for a crime already committed but as a price potential offenders would have to pay for a future crime.

[1] For example, refer to Cohen (2000) for costs of crime and Viscusi (2000) or Spengler (2004) about the value of (a statistical) life and further references.

[2] In fact, Cesare B. Marchese Beccaria published his work 1764 in Italian and its translation was published much later. Jeremy Bentham's manuscripts were written about 1770 but found and published many years afterwards.

[3] In fact, general deterrence would be more precise. In contrast, specific deterrence relates to the offender and aims at deterring him from further delinquency.

Empirical tests of the effectiveness of deterrence began just “recently” in the 20th century. However, after the formulation of a formal model by Becker (1968) and its empirical verification by Ehrlich (1973) an increasing amount of literature emerged which scrutinized the theory of deterrence and its empirical application.

Deterrence is embedded in a body of theories of understanding crime. While it is based on the idea that people adjust their unlawful behavior to changing incentives - expressed by the probability and severity of punishment - many other theories exist on why people offend; from genetic characteristics to social and cultural differences. We describe a selection of such theories in more detail in subsection 2.1.3. Nevertheless, most of these theories can be encompassed by an economic framework: a crime will be committed if the benefits from it exceed its expected costs. While any exact identification and determination of these abstract measures seem to be impossible, criminal behavior - very often expressed by official crime rates - should change when the probability and severity of punishment changes. Most of the empirical studies exploit this principle and make it subject to statistical tests to find out whether any evidence of a deterrent effect can be found. Literally hundreds of such studies have emerged in the last four decades and have been subject of an intense debate. The discussion has been - and still is - especially heated about the question whether the death penalty deters crime or not. However, for almost all offenses two studies can be found which come to completely different results; one finding strong support for a deterrent effect while the other cannot find any evidence of it being at work. This situation is, at least, very unsatisfying from a scientific point of view. Moreover, public policy would obviously greatly benefit from a better understanding of the effectiveness of deterrence.

The large number of available studies, the heterogeneity of their results, the scope of studied populations, offenses and implemented techniques offer a perfect playground for a quantitative analysis of the literature. While many qualitative literature surveys of studies covering the deterrence issue have been published (see subsection 2.1.2) there are almost no analytical reviews to be found. While Antony and Entorf (2003) and Müller (2003) were first steps and feasibility studies of a meta analysis, we are only aware of Pratt (2004) as one further quantitative literature survey[4]. However, the latter considers deterrence only at the margin and focusses on the differences between several selected theories of crime. To the best of our knowledge, the cross-disciplinary project “Metaanalyse empirischer Abschreckungsstudien - ein quantitiver methodenkritischer Vergleich kriminologischer und ökonomischer Untersuchungen zur negativen Generalprävention”[5] was the first comprehensive attempt to use the existing bulk of studies to identify the driving factors behind the heterogeneity of results and to analytically assess the current situation of research. Some preliminary results - with a snapshot of the acquired data - are published in Dölling et al.

[4]Although not an analytical review, Eide et al. (1994) calculate, using a small set of selected studies, some simple bounds of the published results for several types of crime.

[5]Translated: meta analysis of empirical deterrence studies - a quantitative and critical comparison of criminological and economical studies about negative general deterrence. Involved members were: Dieter Dölling, Dieter Hermann, Andreas Woll, Armando Häring from the University of Heidelberg and Horst Entorf and Thomas Rupp from the Technical University of Darmstadt.

(2006) and Dölling et al. (2007). This thesis originates from this work and utilizes the full data set of all 700 acquired studies. In the following chapters we address several questions in more detail: are there any key factors which determine the results of a study (e.g., the studied population, the statistical methods employed, the cultural background of the authors, the studied offense, etc.)? Is there any significant deterrent effect overall? How reliable is the retrieved information? Since there are almost no prior theories about the strength and direction of any potential key factors we resort to methods of data mining. Afterwards, we employ several tests to assess the quality of the calculated estimators, i.e., how well the estimators perform in reproducing and forecasting results.

Indeed, we can identify several elements of the design of a study, the cultural background of an author and offense-specific properties which affect the outcome of a study. Whether or not these elements measure a direct effect or - to some extent - pick up other neglected effects belongs to the subsequent interpretation. Nonetheless, our results should contribute to the knowledge of crime and the understanding of its literature.

The thesis is organized as follows. Chapter 2 summarizes the theoretical background of several theories about deterrence and - with a focus on the rational choice approach - the corresponding problems and empirical and statistical issues. It also shows the large variety of fields the theory of deterrence is applied in and that a lot of contradictory results exist in the literature. The large body of inconsistent results is one of the main reasons why a meta analysis should be helpful to increase the understanding of deterrence. Chapter 3 contains the creation of the data base and its statistical analysis. Several techniques are used to identify important factors which may determine the results of an individual study. Chapter 4 then puts these results into perspective and shows how reliable, trustworthy and usable these estimates are. Subsequently, the results of the best models, in regard to precision and fit, are discussed in more detail. Finally, chapter 5 concludes this thesis and recapitulates the main issues and results. Furthermore, the appendix contains some minor findings which are interesting but not essential as well as an extensive description of all available variables and displays all included studies (accompanied with some important additional information).

2 Deterrence and Crime

If people are good only because they fear punishment, and hope for reward, then we are a sorry lot indeed.

Albert Einstein

Although many other facets of crime and deterrence are considered, this work concentrates on the rational choice theory and its accompanying literature. The literature we retrieved and included almost evenly covers the subjects of sociology, criminology, economics and other fields. Nevertheless, most of the sociological and criminological literature was covered by the team in Heidelberg and, as a result, will be somewhat under-represented in this introductory chapter.

2.1 Rational Choice Theory

> *Unlike Marxian analysis, the economic approach I refer to does not assume that individuals are solely motivated by selfishness or material gain. It is a* method *of analysis, not an assumption about particular motivations. [...] The analysis assumes that individuals maximize welfare* as they conceive it, *whether they be selfish, altruistic, loyal, spiteful, or masochistic. Their behavior is forward-looking, and it is also assumed to be consistent over time. In particular, they try as best they can to anticipate the uncertain consequences of their actions. Forward-looking behavior, however, may still be rooted in the past, for the past can exert a long shadow on attitudes and values.*
>
> *Becker (1993)*

Many disciplines have developed different theories to explain criminal activity. The main influential factors "vary from emotional and behavioral characteristics in psychology, physiological characteristics in biology, environmental and organizational variables in sociology, to alternative cost and benefit consideration in economics" (Howsen and Jarrell, 1987). We do not discuss the psychological approach anywhere and defer to psychological text book material[1] such as Bandura (1969) or Schwartz (1984), but concentrate on the legal, criminological and economical theory[2]. Most of this literature emerged after the seminal work from Becker (1968) who renewed old ideas[3] from Marchese Beccaria (1819); Bentham (1830) and Chadwick (1829) and molded them

[1] According to Gneezy and Rustichini (2000), the first psychological study to test the effect of punishment was that of Estes (1944).

[2] Although we exclude the psychological approach from our analysis we include several studies from psychologists or psychological journals.

[3] Adam Smith had already written about crime and economics.

into a modern theoretical and economic framework (Bodman and Maultby, 1997). In a broader sense, rational choice theory[4] assumes that people, criminal or not, respond to incentives. Each individual may have a different inherited "taste of crime", but as the costs of crime change, the individuals choice is also likely to change. To optimize welfare certain levels of probability and severity of punishment have to be chosen while tolerating a "natural" crime rate[5]. This boils down to the question "whether crime income is positively related to the perceived risk of crime" (Viscusi, 1986). Although Becker's reasoning is obvious, if not banal (Blaug, 1980), it took a long time until science started to analyze this concept (theoretically and empirically). This economic view of crime is fairly compatible with the sociological point of view, given in Erickson et al. (1977), that "the doctrine reduces to the assertion that when a criminal act is contemplated the *perception* of a high risk of a swift and severe reaction by legal officials is a *sufficient* condition for omitting that act". The economic distribution to deterrence research has been acknowledged by criminologists and sociologists:

> *Economists should be given credit for one of the most exciting developments in research on deterrence in recent years. The techniques they used, the controversy they created and the discussions they generated have stimulated interest in deterrence research beyond all expectations.*
>
> *Fattah (1983)*

In general, a model to describe the behavior of potential offenders includes, besides many social, economical and environmental covariates, several deterrence variables. These should describe the actual mental process in the decision whether a crime shall be committed or not. Regarding deterrence the probability that the crime is detected, cleared, the offender arrested, the offender prosecuted, convicted and punished are relevant. Additionally, the severity and type of punishment itself (e.g., a fine, probation or imprisonment) - if a punishment occurs - is also important[6]. The model of Becker implies that, if one of these probabilities and severities increase, the probability of committing a crime will decrease. Introducing more detailed levels of punishment can alter this mechanism slightly.

Typical models are based on the notion that each individual has to allocate his time between legal and illegal activities[7] (Ehrlich, 1973). The potential offender will then commit a crime if his expected utility from legal alternatives is smaller than the expected gain[8] from his illegal actions. Whereas "*Utility* is nothing more than an economist's jargon to represent the personal satisfaction one receives from various pecuniary and nonpecuniary pleasures in life" (Cohen and Simpson,

[4]Rationality should not be confused with narrow materialism (Williams and Sickles, 2002).

[5]Below a certain level of crime the costs of increasing the probability and severity of crime surpass the marginal loss of welfare due to crime.

[6]In principal, these variables are not independent. Punishment, for example, may be less effective when its probability is very low.

[7]Ehrlich (1972) remarks that if an individual is solely active in the illegal sector his actions are inelastic to small changes in deterrence.

[8]The expectation includes - in theory - all possible consequences.

1997). "In the field of legal studies the deterrence theory justifies punishment as deterring future crimes on the assumption that a higher expected punishment produces lower levels of criminal behavior." (Gneezy and Rustichini, 2000). These behavioral implications are then aggregated from the individual level to the entire society.

Of course fundamental criticism followed. Fattah (1983), like many others, argues that criminals are not rational in their actions and are more likely to act because they see no possibilities in the legal sector. The value of monetary gain (and the unpleasantness of punishment) "varies greatly from individual to individual depending on a host of factors" and may, as well as the inaccurate knowledge of the population about deterrence measures, impede statistical analysis. An even more radical argument is "that economic analysis lacks the conceptual resources needed to criticize intrinsic value retributivism[9]" (Kahan, 2004). Furthermore, the perceived probability and severity of punishment is decisive and the correlation with official statistics may not be as good as they are supposed to be (Chaiken et al., 1974). These statistics may inhibit so many errors, inaccuracies or high spatial and temporal sensitivity (including unknown lag structures, differently affected groups, etc.) that cannot be overcome by statistical techniques, or that any results are too fragile to draw any relevant policy conclusions (Decker and Kohfeld, 1990a). For public policy decisions it may not be sufficient to know that deterrence is working in general when the precise effectiveness is unknown (Fattah, 1983). However, some authors like Virén (1994) argue that it is the lack of "good" data which is responsible for mixed or insignificant results. Although large statistical improvements have been made in the last decades, some problems still remain, as will be shown further on. Fattah (1983) reminds us that methodological complexity does not ensure good results: "the sophisticated techniques they use may create illusions about the accuracy of their findings and may give the impression that the results are as good as the techniques themselves".

However, it is important to know that the deterrent effect is unlikely to be proved or disproved on theoretic grounds. When the restriction of fixed leisure time is removed some odd implications become possible: e.g., the normality of illegal activities or the independence of decisions in the legal labor market and variations of the parameters in the illegal market (Heineke, 1978a). Furthermore, Block and Heineke (1975) show that when the time allocation is explicitly entered in the utility function no unambiguously static results can be drawn[10]. Since the effectiveness of deterrence cannot be proven theoretically it remains an empirical matter to verify its impact (Brier and Fienberg, 1980). The latter may be the single issue in the deterrence literature which is almost uncontroversial.

The first study to reach a broad audience whilst being based on the Becker model was conducted by Ehrlich (1973), although Ehrlich (1972) was published one year before. He uses cross sectional data of U.S. states from 1940, 1950 and 1960 and Index I crimes. Using OLS, 2SLS and SUR

[9] This refers to the backward-looking intuition that punishment should be in proportion to the reprehensibility of the committed crime.

[10] The sentence length has a negative substitution and a positive income effect. Strong restrictions are necessary to achieve unambiguous results.

estimators he concludes - among other results - that all crimes vary inversely with the probability and severity of punishment.

An intense discussion immediately followed. Vandaele (1978b) corrects Ehrlich's data for some errors and repeats his analysis with and without outliers and some different specifications. Overall, he confirms the deterrent effects found by Ehrlich. Using this data, Raftery et al. (1997) rely on Bayesian Model Averaging (which should recognize the potential absence of relationships better than other methods) to select the relevant variables and conclude that only the probability of punishment deters but the severity does not. Pogue (1986) argues that most of the significant deterrents arose due neglecting other important variables (see also section 2.3.1). He concludes that there is, using Ehrlich's data, only a small deterrent effect for robbery to be found, if at all. Many other researcher reevaluated the deterrence hypothesis with Ehrlich's data (like Nott and Green (2004); Fernández et al. (2001a); Andreoni (1995)) and come to different conclusions - some offer full support of the deterrence hypothesis, while some find no effect. Brier and Fienberg (1980) reject the Becker-Ehrlich model itself and discard all subsequent empirical findings. They argue that the model is too flawed and the empirical findings (in the economic literature) are not reliable enough to draw any conclusion (neither for or against the deterrence hypothesis). Although their criticism seems to be exaggerated the ambiguity of findings in the deterrence literature was and still is perhaps the main reason for the prolonged debate. Virén (2001) proposes allowing criminal activities to be part of the leisure activities of each individual. He uses mostly Finnish data and concludes, besides finding significant deterrent effects, "that crime also depends on the available amount of leisure time at least to the extent that we are dealing with part-time criminals".

"Becker's assertion that maximization of social welfare requires the exclusive use of fines whenever they are feasible" (Ehrlich, 1981) does not hold when incapacitation effects are taken into account. Ehrlich argues that "even when feasible, fines should be replaced by, or used in conjunction with, an incapacitating penalty" to reach more optimal results. Contrary to the public point of view that fines should increase with recidivism, Emons (2003) argues that, according to game theory, it is optimal to confiscate the whole wealth of the delinquent and none for recidivism (since no wealth remains). However, Garoupa (1999) argue that high fines are not optimal anymore for lesser offenses when there is uncertainty about the probability and severity of punishment. Furthermore, Rubinstein (1980) shows that for every two-level game with a maximum penalty there exists a lesser penalty and a utility function so that the lesser penalty deters more than the maximum penalty. When relaxing the perfect information assumption, Levitt (1997a) shows that "private information greatly reduces the usefulness of fines due to the additional incentive-compatibility constraint that binds the social planner". Another reasoning is proposed by Garoupa and Jellal (2002) who ponder the possibility of collusion between criminals and enforcers. They argue that higher penalties are linked with more resources spent on the detection and punishment of corruption (since the likelihood of collusion increases). "Thus, the government could reduce this sanction [for the underlying offense], save on detection, and increase the criminal sanction for corruption (in order to offset the negative effect on deterrence)". Fines and probations are, as is generally assumed, less threatening than imprisonment. Therefore, increasing the fine- or probation rate leads to a decrease in the imprisonment rate and can therefore lead to more crime due to a reduced level of punishment (Entorf, 2003). However, some studies, like Wolpin (1980), find a deterring effect of the probation rate.

Inspired by German penal law, which relies heavily on fines, Cherry (2001) studies the impact of fines on Index I crimes in California. He finds that the probability of punishment and the fines are significantly deterring crimes while the average sentence length is not. He also points out that an increased usage of fines for non-serious offenses would make the legal system of the U.S. more efficient and less expensive because these offenses make up for the largest part of the U.S. prison population. While the U.S. has (with Russia) built up the largest (and most expensive) prison system in the world, Germany has (by increasing the usage of fines) lowered such expenditures while holding crime constant. He also reports that recidivism is lower for those punished by fines than by imprisonment. However, Withers (1984) notes that, depending on the severity of crime, imprisonment might pay off when the incapacitation effect is large enough.

If an offender is risk neutral the probability and severity of sanction are of equal effectiveness. If he is not - although the expected gain (and thus expected punishment) remains the same - his expected utility changes differently. Becker (1968) shows in his model that an individual who has a preference for risk is more deterred by an increase in the probability of punishment than by a comparable increase in the severity of punishment. The contrary is true for risk-averse individuals. It is even possible that a risk preferrer engages more in crime when the average punishment increases. This may happen if the stimulation effect to commit more offenses due to his reduced expected wealth is larger than the opposite substitution (legal for illegal activities) effect (Ehrlich, 1972). In empirical studies this is used to interpret the behavior of the studied subjects. If the elasticity of the deterrent effect of the probability of punishment is larger (smaller) than that of its severity the studied population may be assumed to be risk preferring (averse). The individual discount rate also influences the risk aversion (Polinsky and Shavell, 1999). For example, an offender with a large discount rate will be deterred more effectively by the probability of imprisonment than by the sentence length.

Many studies find that the probability of detection deters more than the severity of punishment, which is interpreted by Kau and Rubin (1975) as evidence that society spends enough on detecting crime such that only risk-preferring individuals engage in criminal activities. They find a deterrent effect of the conviction rate but not of the average time served and argue that this implies that the remaining potential offenders in the crime market are risk-preferrers. Mendes (2004) argues that, on an aggregated level, "potential criminals mentally combine the three deterrence components[11] - regardless of whether they are risk neutral, averse, or acceptant". However, not all offenders are completely neutral towards the expected gains. Shachmurove et al. (2001) study burglaries in Greenwhich and conclude that burglars prefer lower risk above more loot.

[11]This refers to the probability of arrest and conviction and the severity of punishment.

Deterrence relies heavily on the available police resources and efficiency. Using U.S. state level data, Cameron (1991) analyzes the police output function and finds that the police are largely working (marginally) inefficiently. Bodman and Maultby (1997) add that, ceteris paribus, police efficiency decreases significantly with rising crime, especially when the system is already operating at its limit (Chambouleyron and Willington, 1998). Additionally, the correlation between police expenditures and police output (in most cases arrests or convictions) is low when the police administrators confuse output maximization with budget or utility maximization. Kau and Rubin (1975) point out that law enforcement might be more effective in combating violent crime - at the margin - so that switching resources from property crime to crime against the person might reduce total crime. However, the average citizen fears property crime more than violent crime in regard to the demand for police (Chapman, 1976).

Usually crime is more prominent in cities which might be partially be explained by smaller arrest probabilities and other factors. Glaeser and Sacerdote (1999) report, using previously found elasticities and data from the National Crime Victimization Survey (NCVS), National Longitudinal Survey of Youth (NLSY) and Uniform Crime Report (UCR), that up to 20% can be explained by the lower arrest rates and 33 - 50 percent by female headed households.

Although Sickles et al. (1979); Witte and Schmidt (1977) and Witte (1980) relate more to specific deterrence, we mention them here (and include them in this meta analysis) because these are pioneering studies using individual data (of prisoners from North Carolina) to test deterrent effects[12]. Sickles et al. (1979) regress the sentence length after release from prison on, among others, the number of previous arrests and the first wage. They find that the first wage influences the criminal career but the future criminal career does not affect the first wage. Previous arrests are correctly signed but not significant. Witte and Schmidt (1977) and Witte (1980) find that both the probability and severity of punishment deter, while the marginal effect of the latter is smaller. Subsequently, Tauchen et al. (1994) and Williams and Sickles (2002) analyze a birth cohort from Philadelphia and use police resources per offense as a measure for general deterrence. The former find that police do deter crime, but do this more effectively for people with clean criminal records, whilst the latter do not.

Most authors use data from the United States, Canada, the United Kingdom, Germany and Scandinavian countries. While a group around Pablo Fajnzylber concentrates on South America, other countries are studied only in very rare cases. Wolpin (1980) studies England, Japan and California and finds that Japan's inherent crime rate is even higher than that of the USA (represented by California). He notes that countries with a high inherent robbery rate have higher clearance rates, lower conviction rates and harsher punishments. Mui and Ali (1997) study crime in Honk Kong and find that, contrary to western nations, unemployment, poverty and foreigners are not associated with crime. However, unemployment (due to the shadow economy) and foreigners (illegal immigrants are expelled immediately) might not have been adequately accounted for. An-

[12] These can be interpreted as studies of general deterrence with individual data; but only prisoner-data is available.

other study (Tao, 2004) uses data from Taiwan (which is, for political reasons, assigned to China), Meera and Jayakumar (1995) use Malaysian data and their overcrowded prisons. Argentina is an exception and is subject to several studies; see for example Cerro and Meloni (2000); Chambouleyron and Willington (1998); Kessler and Molinari (1997) or Balbo and Posadas (1998). In some cases multiple countries all over the world are studied (as pooled cross sections or panel data). Usually homicide and robbery are analyzed because these crimes are defined very similarly across all nations. Fajnzylber et al. (2002b) and Fajnzylber et al. (1998) report that economic growth and inequality are main influences of the (anti-cyclical) movements of crime. Western Europe stands out by having high homicide rates, South America by robbery and Hindu countries by having very low homicide and robbery rates. Recessions and other shocks can lead to persistent long-term increases in crime.

Assessing the available literature, the severity of punishment is generally assumed or found to be less relevant than the probability of detection or punishment. An exception to this rule is the study by Funk and Kugler (2003b), who focus on this issue and conclude, using data from Switzerland, that both - probability and severity of punishment - are of equal importance. Another aspect is the interaction of the probability and severity of punishment, especially when the probabilities are very small. A few authors explicitly test the expected punishment and come to different results: Curti (1999) finds significant deterrent effects using German data, Swimmer (1974a) only finds deterrent effects for aggravated assault and burglary (and larceny in a non-linear estimation) using data from all U.S. cities with more than 100000 inhabitants. However, probability and severity of punishments are not always substitutes. As described in subsection 2.2.4, this is found in an experiment with students about free riding (Anderson and Stafford, 2003).

It is generally assumed that the introduction of harsher maximum penalties, mandatory penalties or larger police budgets should deter potential offenders. These effects may be mitigated for several reasons. Andreoni (1995) points out that, while more punishment has a deterrent effect, criminals will invest more energy on avoiding capture. Furthermore, judges and juries may be more hesitant to convict someone as punishment increases (Bodman and Maultby, 1997; Vingilis et al., 1988). More severe maximum penalties cannot be effective if the expanded scope of sentencing is not utilized. For example, in 1991 new sentencing guidelines were introduced which were aimed at doubling the median fine for corporate crimes. However, Parker and Atkins (1999) do not find a significant change in the imposed sentences. Consequently, they can not find much evidence for any increased compliance. Stafford (2002) uses an increase of almost 2000% for fines (violating waste regulations) but finds only very small effects. Increasing mandatory minimum sentences avoids this problem (at least for the distribution of the lenient sentences) but may lower the probability of conviction. A judge or jury may be more cautious convicting someone when the minimum sentence exceeds the penalty they would have imposed otherwise. When these effects, which lessen the deterrent effect, are incorporated into the deterrence model, former significant results become insignificant. This is supported by Mustard (2003) who reports that judges in Oregon lessened their overall sentences when minimum sentences were introduced.

Witte (1980) additionally points out that social conditions also modify the probability of punishment. Using individual data, being married or having a job did influence the conviction probability but not the arrest probability. Moreover, many studies use the police budget as a proxy of police effectiveness. Although it is quite reasonable to assume that a larger budget enhances the available equipment and training, or increases the available manpower, there may be some mitigating effects. On the one hand the police may be interested in high crime rates to avoid decreasing budgets (see also section 2.3.1 for this feedback effect). Furthermore, the police may redeploy their forces to crimes which, overall, reduces the marginal deterrent effect of the police budget or simply to crimes which are not considered in the usual analysis (see subsection 2.2.3 for an example).

The focus of the empirical studies has changed over the last decades. In the beginning, researcher were mainly interested in the deterrent effects of probabilities, sentences and laws. However, "the economics literature on crime has transited from an emphasis on economic conditions (including education) and deterrence effects to more recent considerations of factors that may explain how crime is propagated over time and within communities" (Fajnzylber et al., 1998). Nowadays, a large part of the literature implements deterrence variables as covariates in studies of other effects (for example, unemployment and crime).

2.1.1 Public Perception

> *In a way general deterrence is a kind of belief. It has been introduced in penal law not after series of investigations in which its validity has been proven. It has been accepted as a useful concept in penal law because people believed in the deterrent influence of sanctions.*
>
> *Buikhuisen (1974)*

As Fattah (1983) points out, deterrence is only applicable to those who are not lawful or criminal *by nature* - or stated economically: "the deterrence effects should be strongest in the group where the expected costs and benefits of noncompliance are closest to being identical" (Braithwaite and Makkai, 1991). This leads to the consideration of the general normative development of the society or sub-groups in the analysis of deterrence. This is often done in studies dealing with driving under the influence, considering those who never drink alcoholic beverages and hardcore drinkers (see subsection 2.2.6) but seldom in other fields. As an exception, for example, Salem and Bowers (1970) study deterrence in regard to minor offenses in U.S. colleges and universities. These offenses are significantly reduced by increasing penalties but deterrent effects are largely rendered irrelevant when the normative attitude of the students are considered. Gertz and Gould (1995) asked 611 college students in Florida about their past delinquency and find only insignificant support of the deterrence theory but judge moral attitudes to be relevant. Indeed, Bohnet and Cooter (2001) point out that laws made to reduce offenses do not only deter but do also educate and coordinate while Cloninger (1994) applies the same argument to police presence. Erickson et al. (1977) argue that “all purported evidence of general deterrence is suspect” until the deterrent effects hold “independently of the social condemnation of crime”. Furthermore, low crime rates in combination with a high perceived certainty of punishment may not reflect deterrence but the social (extralegal) attitude (“what ought to be”) toward these offenses. Brier and Fienberg (1980) also argue that any educational effect of deterrence might not be distinguishable from the pure deterrent effect. As pointed out in subsection 2.2.6, this may be seen in the case of drunken driving when the probability and severity of punishment and the public awareness of the problem rose simultaneously.

Police crackdowns in crime prone areas have been implemented into the resource distribution strategy of the police, especially for drug-related crimes (Sherman and Rogan, 1995) and Driving Under the Influence (DUI), as pointed out by Benson et al. (2000). Besides increasing the efficiency of police actions these crackdowns are also likely to increase the public awareness and acknowledgement of police activities. Furthermore, the private sector also responds to crime (Clotfelter, 1978) by investing huge amounts of money in security measures (Witte and Witt, 2001). However, Guttel and Medina (2007) show in a game-theoretic model that “such investment will not only affect the behavior of the perpetrators, but will also affect that of the police”. In the case that the police concentrate more on the protection of the more vulnerable sector this may impede spending on private protection measures in that sector. Furthermore, beside the usual security measures like alarm systems, better locks and other equipment, keeping weapons at home and, especially, carrying concealed guns may even increase crime (refer to section 2.2.9 for a more detailed description of the corresponding discussion).

Deterrence research is of great societal importance. Fear of crime and policing against it is a major topic in many societies and scientific disciplines. Besides criminology, sociology, economics or public policy, even the CDC (Centers of Disease Control and Prevention) and the WHO (World Health Organization) are interested in deterrence research. The latter have declared violence prevention a public health priority: where criminal justice emphasizes punishment, deterrence and incapacitation, public health focuses on primary prevention (Foege et al., 1995). Although advises for public policy may be given from theoretical and empirical studies, it cannot be taken for granted that these are incorporated into actual laws accordingly. In the famous case[13] Gregg v. Georgia which ended the death penalty moratorium in the U.S., Justice Stewart (U.S. Supreme Court, 1976) wrote about the deterrence hypothesis of capital punishment that “there is no convincing empirical evidence either supporting or refuting this view”[14]. Sampson and Raudenbusch (1999) state that the broken windows theory (introduced by Wilson and Kelling (1982); see section 2.1.3) was the determining factor in the police crackdowns in several cities (especially New York). Also, Weber and Crew (2000) note that the laws to ensure water pollution control were mo-

[13] Solicitor General Robert Bork introduced Ehrlich (1975a) - who found strong deterrent effects of capital punishment - into the case.

[14] Nevertheless, they were convinced that there exist at least some potential offenders who are deterred but it remained nebulous on what their opinion was based on.

tivated directly by deterrence research. However, due to the great variance of methodologies, data sources and results, great care has to be taken when deriving any public policy advises: "a cynical view would point out that familiarizing the public with the research findings would give people the information to distinguish between political promises about crime control that are merely wishful thinking and promises that might have merit" (Becsi, 1999). Daily political business is not likely to mitigate this dilemma because, as can be verified on a regular basis, crime statistics and isolated studies are not appropriately put into perspective. The following quote illustrates the point of this problem:

> *The release of new crime statistics is typically followed by a barrage of partisan political approvals and disclaimers depending on which party or interest group benefits. Incumbents are always quick to accept the credit for any decrease in crime rates while opponents are just as quick to challenge the reliability of the statistics or argue rates would have somehow fallen faster had they been in office.*
>
> *Doyle et al. (1999)*

2.1.2 Literature Surveys

> *Read not to contradict and confute, nor to believe and take for granted, nor to find talk and discourse, but to weigh and consider.*
>
> *Francis Bacon, 1625*

In the last five decades several surveys have been published. Among these we consider Cook (1977); Nagin (1978); Brier and Fienberg (1980); Beyleveld (1980); Cameron (1988) and Eide et al. (1994) to be the most prominent.

In an early review Cook (1977) comes up with a mixed conclusion. While "there is strong evidence from some of these studies that an increase in the threat of punishment can reduce the amount of some crimes in some circumstances" he attenuates this by remarking that these evidence do not cover long-term effects, magnitudes involved and other points. He thinks that it is "highly unlikely that anything like a complete scientific basis for criminal justice policy will be produced in the foreseeable future".

After Becker's and Ehrlich's initial work Nagin (1978) is the first popular survey of the deterrence literature which was in its infancy at that point of time. He evaluates about 20 studies and delves into several sources of possible problems (data acquirement on the police level, simultaneity and identification issues, incapacitation effect, etc.). He estimates that about 20 to 80 percent of the decline in crime could be attributed to incapacitation. He ends with the insight that it is too soon to draw any final conclusions but deterrence seems to be working. He concludes that "a more critical assessment of the evidence is needed if we are to see progress in the development of knowledge about deterrent effectiveness and its application to effective public policy".

Brier and Fienberg (1980) draw very pessimistic conclusions. They conclude that no progress had been made in ten years of research and much of the debate about Ehrlich's findings only diverted the efforts of serious scholars. Moreover, they also state that no reliable conclusion can be drawn about the deterrence hypothesis and they “believe that little will come from further attempts to model the effects of punishment on crime” with that kind of employed data.

Beyleveld (1980) gives a profound summary of the literature about general deterrence from 1946 to 1979. After giving a good and generalized introduction into the deterrence theory he presents reviews (some very short, some very detailed) and comments on 110 theoretic and 216 empirical studies. He also discusses many of the different approaches and problems of analysis in empirical studies.

Cameron (1988) first portrays several different and important theoretical aspects of the deterrence theory. Then he surveys the empirical evidence of 79 studies and especially comments on methodology issues of these studies (e.g., identification problems, variable structures, etc.) and on their contradicting results. As the studies above, he sums the existing literature only qualitatively and gives a subjective picture of the current situation at that time. He concludes that “much of the literature seems impaired by bias due to measurement error” and argues that there is need for a further development of the underlying theory.

In their book Eide et al. (1994) give a thorough overview of the determinants of crime, the rational offender and utility-based models. They then delve into the methodological issues of empirical studies based on macro data. They review shortly 15 correlational studies and then report the results and properties of 21 cross-sectional studies; including the socioeconomic coefficients of these studies. In contrast to the surveys above, they give some statistics of the studies’ results: a table with the bounds of the retrieved elasticities and their corresponding median for various methods. This can be interpreted as a first small step towards a numerical summary of the deterrence literature. Furthermore, they review 18 other empirical studies using time series and individual data. The book is concluded with several analyses of Norwegian data with conventional and more sophisticated methods.

Even after more than 30 years of empirical research Polinsky and Shavell (2000) still remark that “empirical work on law enforcement is strongly needed to better measure the deterrent effect of sanctions, especially to separate the influence of the magnitude of sanctions from their probability of application”.

2.1.3 Other Theories

Note that the ecologist would say that the environment is causing the criminal to act, whereas the economist would say that the criminal is acting taking his environment into account.

Chapman (1976)

This quote puts it aptly that there is more than one way to understand criminal behavior - even if the reasoning is fully compatible with rational choice theory. While many studies implement covariates which could be interpreted in the setting of the following theories, most authors do not comment on them in this fashion. Refer to the meta analysis done by Pratt (2004) for empirical studies which specialize on some of the subsequent theories. These approaches should be interpreted as complements and not as competing theories. Some summaries of the following theories are condensed versions of the descriptions given by Pratt (2004).

Broken Windows Theory

This theory, introduced by Wilson and Kelling (1982), suggests that individuals more readily engage in crime if their neighborhood shows signs of decay. In this scenario broken windows, abandoned buildings, graffiti or simply very lenient treatment of misdemeanors imply a lower (perceived) probability that an offense will be prosecuted or cleared. However, social decay "is accompanied by physical deterioration, as homeowners and small business people put less time and money into maintaining their buildings"(Spelman, 1993). Therefore, broken windows may be just an indicator for the underlying social decay which is accompanied by more crime.

Lochner (2001) tests the broken windows hypothesis and finds no effect. Perceived arrest probabilities (for theft and burglary) are uncorrelated with various neighborhood characteristics (gang-activity, lawlessness, abandoned buildings or drunks on the street). In another study he finds no relationship between the beliefs of the probability of arrest and the information about the arrests of other random individuals and local neighborhood conditions (Lochner, 2003).

Corman and Mocan (2002) test the hypothesis in New York by including the arrest rates of misdemeanors among the crime specific arrest rates, police manpower and prison population for various Index I offenses. Significant effects are only found for robbery and motor vehicle theft. Kelling and Sousa (2001) also use New York data but only the arrest rate for misdemeanors is negative and significant. They interpret this as an approvement of the broken windows theory.

Funk and Kugler (2003a) approach the broken windows theory by analyzing the effect of lesser offenses (burglary and theft) on more severe crime (robbery). Using Swiss data they find that an increase in burglary and theft lead to an increase in robbery but not conversely. They conclude that a tougher enforcement on minor crimes also reduces major crimes.

However, concentrating on minor offenses may also backfire because arrest for misdemeanors may later increase severe crimes (Sherman et al., 1998) and resources tend to be misallocated. For example, neighborhood watches are not used where it would make sense (lack of trust) and are used where it does not make much sense (neighborhoods in the middle class with low crime rates). Contrarily, they point out that such neighborhood watches may increase the fear of crime while having no measurable effect on crime rates. Nonetheless, having good ties to the police reduces crime significantly.

Similar to the arguments of the broken windows theory, Posada (1994) theorizes that when random increases in crime are not counteracted by more efforts in deterrence, the perceived rate of apprehension decreases and the random increase in crime may become permanent. This leads to the notion of Sah (1991) that "past crime breeds future crime".

Life-Cycle Theory

Basically the theory states that criminals start their career by committing minor crimes and then, by accumulating criminal capital, proceed to more severe offenses. Usually, (detected) criminal activity then recedes when a certain age is reached (see also subsection 2.3.4). Therefore, the age structure affects crime trends. However, Marvell and Moody (1991) argue "that the age/crime relationship is probably exaggerated because the high arrest rates for younger persons are due partly to their lesser ability to escape arrest, younger persons commit more group crime, and the age structure of victims should be taken into account".

Funk and Kugler (2003a) present evidence, using Swiss data, for the increasing severeness of committed offenses as the criminal becomes more experienced. Marvell and Moody (1991) analyze 90 studies about the relationship between the age structure and crime. Although there is only little evidence they cannot conclude that there is no relationship. However, forecasts based on demographic trends do not seem to be helpful in explaining crime.

Anomie/Strain- and Social Disorganization Theory

The social disorganization theory goes back to Shaw and McKay (1972) who observed that juvenile crime was not distributed evenly in Chicago but was concentrated on "slum neighborhoods" regardless of the local ethnic composition. The observation of receding individual crime rates when moving to less crime prone areas lead to the conclusion that crime is a function of neighborhood dynamics and not necessarily of individual characteristics. Low social ties, high mobility, low socioeconomic status and "criminal traditions" are characteristics of these "slum neighborhoods". The theory was very popular in the 50s and 60s until interest shifted from group dynamics to individual processes and has now become more popular again. This recurring interest is - at least to some extent - based on the inclusion of "intervening mechanisms"; the indirect effect of social disorganization on crime via other variables (like family disruption). Kelly (2000) finds that economic factors and deterrence seem to be important for property crime, while social influences, in line with the strain and disorganization theory, are better suited to explain violent crime.

The anomie (or strain) theory was developed by Merton (1938) who argues that the rigid adherence to conventional (American) values may also foster crime whereas the disorganization theory is based on the rejection of these values. In the USA more emphasis is put on (visible) economic success and the pursuit of the "American Dream" (that working hard enough will eventually pay off). However, poor people are more limited in their possibilities in this race to success. This difference between compulsion and limited possibilities lead to a weakening of cultural norms which Merton calls "anomie". As with the social disorganization theory, interest in the the anomie theory was very high in the 50s and 60s and then receded until it rose again in the 90s when Messner and Rosenfeld (1997) reformulated the theory and integrated "an institutional structure dominated by the economy" (Pratt, 2004).

Absolute Deprivation/Conflict Theory

Conflict theory (Bonger, 1916), in a simplified form, divides society into two groups with the upper class having more political authority or social power than the lower class. Crime is then interpreted as a label put on some behavior of the lower class by members of the upper class. As a consequence - although the same deviant deeds are committed by members of both classes - prosecution and punishment tends to be more intense for the lower class. While some areas seem to fit this theory quite well - like conflicts between workers and management or victimless crimes like vagrancy - classic crimes like robbery and rape are more problematic.

Absolute deprivation theory essentially gives some explanations where crimes may originate from. First, poverty - as an important characteristic of the lower class - may directly cause crime. For example, theft and robbery may be (at least subjectively) necessary for some people to survive. Second, poverty may be viewed as a consequence of a "wrong" social arrangement and thus indirectly cause crime when the lower class strives for a change.

Relative Deprivation and Inequality Theory

While absolute deprivation theory is based on the absolute poverty of social groups, Blau and Blau (1982) point out that "racial socioeconomic inequalities are a major source of much criminal violence" rather than absolute poverty. Although not being poor by monetary standards people may think that society withholds something which they are entitled to. It is the inequality which induces subcultures (especially for youths) which may "bring young persons into contact with the law". Furthermore, they observe that "aggressive acts of violence seem to result not so much from lack of advantages as from being taken advantage of, not from absolute but from relative deprivation".

Routine Activities Theory

Cohen and Felson (1979) identify three kinds of measures relevant for crime: motivated offenders, suitable targets and absence of capable guardianship. "Convergence in space and time of the three minimal elements of direct-contact predatory violations" is assumed to be correlated with increased victimization. They define "routine activities" as the "recurrent and prevalent activities which provide for basic population and individual needs". Daily routines and - to some extent - economic success increase the amount of potential victims and reduce the presence of guardians. The theory presumes the existence of motivated offenders and provides little means how to change these motivation. As Pratt (2004) puts it: "presented with opportunities (suitable targets) divorced from capable guardians (either formal or informal), *crime happens*." Cohen et al. (1980) present a variation of the routine activities approach by concentrating more on the situational opportunities in a given area.

Social Support and Altruism Theory

Social support or altruism theory focusses on the relationship between characteristics of social aggregates and the insulation of crime. Social support is viewed as provisions supplied by the community, the government, social networks, the family and other sources (Cullen, 1994). Social altruism is also assumed to be negatively correlated with crime rates. Chamlin and Cochran (1997) define social altruism to be "the willingness of communities to commit scarce resources to the aid and comfort of their members, distinct from the beneficence of the state". Sometimes it may be unclear whether a state is supportive or altruistic because activities (like rehabilitation) can be attributed to both areas. Furthermore, any crime reducing effect may belong to other side effects (e.g., better socioeconomic conditions resulting from a rehabilitation program).

Subcultural Theories

Following the social support and altruism theory, social and cultural conditions may prevent people from engaging in crimes. Hence, it is assumed that there exist social and cultural influences which increase crime. Sources of such violent or deviant subcultures are presumed to be found in urban areas and - only applicable for the United States - in the South.

In the 20th century the South of the United States stood out with its high rates of violent crime. Some researcher argue that "certain cultural norms contained in the South may predispose individuals to not only engage in violent behavior, but also to approve of such actions on the part of others" (Pratt, 2004). Many different cultural norms are identified to be such factors like the historical tradition of chivalry or the tendency to resort to violence when defending the honor of a woman. Many explanations are given why these norms are concentrated in the South (e.g., different religious perspective or the bitterness of having lost the civil war). Also other factors are characteristic for the South like the high rate of firearm ownership.

The often found positive relationship between the size of the population and the crime rates have also been subject to a cultural interpretation. Following Fischer (1975) the probability that people with unconventional lifestyles and interests meet each other increases with the population size. Second, urban regions provide the opportunity for people with unconventional lifestyles to form subcultural groups. Finally, these groups compete against each other for geographic and social space. This leads to a greater identification with group-values and larger within-group cohesion. And, "since subcultural values tend to follow a process of *diffusion* from one generation to the next, the positive association between population size and rates of crime and deviance will tend to persist over time" (Pratt, 2004).

Reintegrative Shaming and Stigmatization Theory

Braithwaite's theory of reintegrative shaming (Braithwaite, 1989) describes punishment as a tool which can either amplify or dampen crime, depending on how it is applied and recognized. Stigma may, on the one hand, prevent crime by the expected reduction in social (e.g., diminished respect and avoidance by friends, family and neighbors) and human capital (e.g., lowered or missing income). On the other hand, once stigmatized, recidivism may become more likely (Fajnzylber et al., 1998). The latter effect is supported in an analysis by Tittle et al. (2003) whereas the former effect is not. Consequently, they recommend that the shaming theory needs more clarification and refinement "to specify more carefully the conditions under which shaming processes inhibit or enhance criminal probabilities". However, using a sample of residents from Shanghai, Lu et al. (2002) conclude that there is no effect within the family but a significant shaming effect in regard to the residents in the neighborhood.

At this point it is worthwhile to note that "individuals asked to judge undesirable behavior tend to explain their own behavior as a consequence of external environmental factors while attributing the deviant behavior of others to poor moral character" (Gottfredson and Hirschi, 1987). For further references see Tittle et al. (2003).

Other Theories

While the above theories are all forward-looking most penal laws also contain a backward-looking aspect: retribution. This means that punishment is not only assessed by its deterrent effect (on the offender and the public) but also by the severity of the offense (Rawls, 1955). This follows from the simple idea that wrong-doing merits punishment: the offender should suffer according to his guilt and the depravity of his crime. "The intention of providing a deterrent is not a purpose of punishment in retributive theory per se" (Avio, 1987).

In the context of drunk driving, Soper and Thompson (1990) note that deterrence is linked to implementation theory (Edwards, 1980). Whenever means of deterrence require bureaucratic action their effectiveness depends on their implementation. These are the communication and the predispositions of public officials towards that policy, the resources of the public officials as well as the bureaucratic structure of the involved agencies.

A completely different approach is the identification of genetic differences between offending and non-offending people. While this does not mean that anyone is a "born criminal" or not, there are genetic differences which seem to increase unsocial behavior. See, for example, Moffitt (2005) for more information and references.

2.2 Particularities Regarding Offenses

Blind commitment to a theory is not an intellectual virtue: it is an intellectual crime.

Imre Lakatos, Philosophical Papers Volume 1, 1977

In the following, many studies - but not all of them - are cited for a number of reasons. While many authors follow the usual way of regressing the crime rate on some other rates and interpret the coefficients, numerous studies contain more information than that. Some authors use or construct unusual variables, new or alternative ways to circumvent problems, introduce innovative ideas or uncommon interpretations to well known variables. Furthermore, there are some data problems in the deterrence literature which require special consideration, aside from the typical difficulties encountered in many studies using cross sections, time series, experiments or surveys. The next subsections deal with such particularities in regard to the studied offenses, various aspects and features of analysis and other interesting aspects. The mixture of different, complementary and contradicting results is also expressed in these subsections.

2.2.1 Classic Offenses

The simplest schoolboy is now familiar with facts for which Archimedes would have sacrificed his life.

Ernest Renan, Souvenirs D'enfance et de Jeunesse, 1887

Spengler (2004) is a unique German study in respect of the utilized data. Merging different data bases from the German states he is able to consider the full punishment cascade (arrest rate, conviction rate, imprisonment- and probation rate, sentence length and monetary fines, see page 43) for adults, adolescents and youths for various types of offenses. After considering many potential methodological problems, the arrest and conviction rates (but not the punishment) are found to be significant deterrents. The same data is used by Entorf and Spengler (2005) who add that the conviction rate is the most important deterrent and, therefore, the increasing discontinuation of criminal prosecution in Germany is counterproductive. In addition to that, the regional distribution of crimes in municipalities in Baden-Wuerttemberg is studied by Spengler (2004). He finds that the mobility of delinquents is of great importance for theft.

Morris and Tweeten (1971) compare the effectiveness of police manpower in 754 U.S. cities. In a 2SLS analysis they estimate the coefficients for the police of all cities and for nine categories (small to large city). Then the coefficients are adjusted - using the available covariates - in such a way that all cities have the same crime rate. The results show that the police are most effective in middle-sized cities and that police deter in all but large (more than one million inhabitants) cities.

Some authors emphasize that short and long term effects should be distinguished when analyzing crime. Usually error correction models are used in these cases. Diez-Ticio (2000) finds short term deterrent effects of the clearance rate on robbery, burglary and auto vehicle theft. However, in the long run only the deterrent effect for robbery remains.

Corman (1981) employs the difference between the ranks of courts to judge deterrence for property crime. A conviction by a higher court implies longer proceedings and more severe records. While arrest rates, conviction rates and sentence lengths have only small deterrent effects, “court processing does matter to the potential offender, even court outcome held as a constant factor”. Therefore, plea bargaining and other measures which result in more cases resolved at lower court level may increase crime rates.

Some deterrents are rather unorthodox. Ayres and Levitt (1998) study the usage of LoJack devices in cars in U.S. cities. These are GPS-senders built into a car which enable a swift recovery of a stolen car by the police. The usage of such devices in an area does indeed deter motor vehicle theft but purchasing is only benevolent for those car owners without an insurance (not the specific LoJack in a car deters, but the general usage in a large area). Goodman (1997) finds very significant deterrent effects when police officers could use their private cars with a visible mark, identifying them as police officers.

It is obvious that crime induces direct costs[15] (broken inventory, injuries, loss of goods, etc.). However, the indirect costs are also manifold (psychic costs, expenditures for protection, avoidance of dangerous areas, abdication of activities, etc.). For example, Philipson and Posner (1996) report that in the USA 300 billion dollars are spent on private security, which is much more than by the government (100 billion dollars in 1995 as reported by Ayres and Levitt (1998)).

2.2.2 Death Penalty

Any punishment has several goals. They include correction and retribution. There is no correction in the case of the death penalty, only retribution. It is even unclear to whom it applies, since a man whom the state eliminates doesn't feel anything after that.

Vladimir Putin, Interfax, 2006

The basic idea of the effect of capital punishment is simple: "if rational people fear death more than other punishment, the death penalty should have the greatest deterrent effect" (Ehrlich, 1977a). Some authors have pointed out (Zimring and Hawkins, 1973), however, that execution might not be the most severe punishment and it is therefore not a question of whether capital punishment deters or not. What matters is the question whether capital punishment does significantly deter more than the alternative punishment (usually lifelong imprisonment). Nonetheless, the focus of the majority of studies lies on the absolute deterrent effect of capital punishment.

Among the 82 studies which do include results on the death penalty, almost all use data from the USA (71) or Canada (5). Avio (1979) is one of the authors using Canadian data. He emphasizes that the Canadian time series (1927-1960) is of better quality, compared to the data Ehrlich (1975a) used. He also points out that "criminal homicide is the relevant variable of interest in an empirical investigation of the deterrent effect, since an offender may not know in advance the legal classification of his crime; the courts alone decide whether the homicide was ultimately a murder or a manslaughter offence". Even resorting to the results which are most in favor of the deterrence hypothesis - using 2SLS and various specifications - no deterrent effect of the death penalty is found. In contrast, Layson (1983) extends Avio's data to 1977 (the abolishment of the capital punishment; but only three executions occurred in that extended time frame) and finds significant

[15] Although a thief may have negative costs because he gains property, these benefits are of minor importance.

deterrent effects of capital punishment. However, he softens his statement in the conclusion and puts more emphasis on the fact that there is an important negative time trend in the homicide rate. A few years later, Avio (1988) published a study in which he puts more attention on the measurement bias of the execution rate. Avio argues that many previous results of studies which use the simple quotient of executions and convictions are overestimated. He uses the proportion of executions to executions and commutations as the probability of execution. He argues that "there is no readily apparent reason why offenders would utilize a less reliable forecast of their prospects", and emphasizes the inferiority of the usual proxies of the risk of execution. Avio (1987) analyzes two models with Canadian data: a retributive model and a deterrence model and finds that both are not significantly different but the latter explains more variance.

One of the most popular studies, which fueled much of the following debate about the effectiveness of the death penalty, was Ehrlich (1975a) and Ehrlich (1977a), who finds very significant deterrent effects of the death penalty. Layson (1985) gives an overview of the ensuing discussion about misspecification problems, data quality, functional forms and other problems. In his study, he confirms the deterrent effect of capital punishment but, contrary to Ehrlich, finds no effect of unemployment or race. Donohue and Wolfers (2005) summarize several studies about the death penalty in the United States, reevaluate them and fix several mistakes. Additionally, they compare the United States with Canada and come to the conclusion that all outcomes are too fragile and that the number of executions is too low to draw any noteworthy and robust conclusion. Using instrumented variables, while considering serial correlation, Ehrlich and Brower (1987) reevaluate Ehrlich's time series data and confirm his previous results. Additionally, they employ a court index of important sentences to represent the court production function and find that it significantly deters homicides.

Hoenack and Weiler (1980) use Ehrlich's (corrected) data and argue that his negative coefficients of the execution rate probably rely on the reaction of the legal system towards the homicide rate. By using a different murder supply function they only find the conviction rate to be a significant deterrent. Passell and Taylor (1977) also criticize Ehrlich's study by arguing that his results are not robust (e.g., under different compositions of variables and transformations). In a reply, Ehrlich (1977b) refutes these arguments, and presents - by using old and new data - new results which support the deterrent effect of the death penalty. Cover and Thistle (1988) also reevaluate Ehrlich's and many subsequent studies and conclude that most of them (all but Layson (1985) depend on the implemented functional form) do not offer stable results and that the deterrent effect of the death penalty has to remain doubtful. Furthermore, the outcome of a study may largely depend on the choice of covariates. Depending on their composition, anything from significant deterrence to absolutely no effect may be found (Leamer, 1983; McManus, 1985; McAleer and Veall, 1989). Cloninger (1977) picks up on much of the criticism of Ehrlich's studies and concludes that executions (and partially the imprisonment probability) are deterrents. Yunker (1976) remarks that old data (UCR data before 1959) is not suitable for studying any deterrent effects (due to identification problems) and many of the usual problems (measurement bias, omitted variable bias, autocorrelation and spurious correlation) can be neglected. He concludes that, using data after 1959, the death penalty significantly deters. Zhiqiang (2004) also reevaluates Ehrlich's data and finds deterrent effects while distinguishing between states with and without death penalty. Abolishment of the death penalty does not only diminish deterrence in that particular state but also in the other states.

The death penalty moratorium in the United States in the seventies is used as a natural experiment by many researchers to study the deterrent effect of the death penalty (Chressanthis, 1989). Almost as much criticism followed these studies. Donohue and Wolfers (2005) emphasize that "the homicide rate in Canada has moved in virtual lockstep with the rate in the United States, while approaches to the death penalty have diverged sharply". While Canada practically stopped executing in 1962, the United States[16] suspended the death penalty from 1972-76. The rise and fall of the homicide rates in Canada was very similar to that in the U.S. although the death penalty remained abolished in Canada. A similar event is used as a quasi-experiment by Cloninger and Marchesini (2001). They use a one-year-moratorium for the death penalty in Texas and find that homicides significantly increased in that year. After executions were resumed the homicide rate receded to its former level although the number of executions were doubled. They argue that this indicates a short-term deterrent effect of the execution rate which diminishes as executions increase. Some crimes do not happen often enough to calculate a rate. Although this basically applies to all crimes and deterrence measures (e.g., robberies in city districts on a daily basis) it may pose a severe problem for executions. Even without a moratorium most countries or states execute people only rarely (compared to other crimes and punishments). To circumvent this problem, some authors assume at least one execution in that period; for example, Ehrlich (1975a) does it to take the logarithm. However, Peck (1976) points out that even very small probabilities may cause problems in such an analysis. Some authors like Avio (1988) take the average of several periods or implement a bayesian belief updating system to calculate execution probabilities for periods without any executions (Layson, 1985).

Levitt (2001) emphasizes that the analysis of panel data is suited better to study the deterrent effect of the death penalty. He uses U.S. state data from 1950 to 1990 and finds no deterrent effect of the execution rate. However, when fixed effects are replaced by interaction terms between state and decade, the imprisonment rate (per capita) is a significant deterrent. He argues that simple fixed effects are not applicable because the state-effects may slowly change over time. Berk (2005) points out that some deterrent results may rely heavily on the data from Texas; Texas executes more people than any other state in the USA. However, in his analysis data from Texas make up only for 1% of all observations. If these are removed (treating them as outliers) any deterrent effect of capital punishment disappears.

Zimmerman (2004) argues that the mere existence of the death penalty does not deter, while

[16] At this point we emphasize that some states abolished the death penalty long ago (Iowa), some states did not execute anyone on death row for a long time (New Jersey executed no one since 1963 and abolished the death penalty 2007) while others do execute fairly often (Texas).

an actual execution does deter within a short time frame (the year the execution takes place). He calls this an "announcement effect", which is reflected by significant execution rates but insignificant conviction rates and is supported by the assumption that potential murderers will be better informed about executions than conviction statistics. Dezhbakhsh and Shepherd (2003) use panel data of all U.S. states and find significant effects - but they did (practically) not consider any measures of the probability of punishment. Dezhbakhsh et al. (2001) do so in another panel data study and find significant effects for all deterrents (arrest-, conviction- and execution rate). However, to be on death row without a pending execution lowers the deterrent effect of the conviction probability for homicide. Both approaches are combined by Shepherd (2004) who uses monthly panel data. The execution rate and the time on death row are found to be significant deterrents. She concludes that an execution has a short-termed deterrent effect - especially in conjunction with a short time on death row. She also emphasizes that many lower homicide rates in states without a death penalty are actually higher when they are adjusted (i.e., standardized by demographic factors). Grogger (1990) uses high frequency data and studies very short termed effects of executions on a daily basis (7-14 days before and after an execution) but cannot detect any deterrent effect. Katz et al. (2003) doubt that a deterrent effect of the death penalty (when it exists) can be shown with the usual data because the variation in the crime rates is simply not large enough: there "simply does not appear to be enough information in the data on capital punishment to reliably estimate a deterrent effect". Moreover, Donohue and Wolfers (2005) note that the debate about the death penalty "may be driven by ideology and advocacy motives". It may even be possible that statistical complexity is used to "silence the debate rather than enlighten policymakers". This is especially important in the case of capital punishment: "unfortunately, the history of the death penalty debate is replete with examples of plausibility being sacrificed on the altar of sophistication".

A completely different financial approach is taken by Cloninger (1992) who interprets the different crime types as assets a criminal can invest in. Therefore, the problem that "those conditions that cause crime rates to rise in general may induce increases in the homicide rate that overwhelm the negative effect produced by executions" may be avoided. He concludes that executions (almost) only affect the homicide rate and deterrence is effective. Instead of the usual UCR data, Sloan et al. (1994) employ data from the National Center for Health Statistics (NCHS) and find that the death penalty and more police are related to less homicides. Avio (1979) notes that in the USA, in 1950, the number of people killed during police operations was three times higher than those executed (while it was only half as large in Canada).

Cloninger and Marchesini (2001) emphasize that it is not a single study which "causes any neutral observer pause" but the sheer number of different studies using different data and methodologies which come to the same conclusion that capital punishment deters. However, Donohue and Wolfers (2005) point out that these "estimates may reflect omitted factors related to the political economy of punishment". A feedback effect may lead to positive (or less negative) correlations because the increased crime rates may increase the demand for more punishment. A contrary effect may be induced when "more vigorous deployment of the death penalty might occur at the same time that the government elects *to get tough on crime*". In this case, all punishments are expected to get harsher and used more often such that the deterrent effect of the death penalty pales in comparison. Furthermore, the public support for the death penalty may depend on the current crime rates: more homicides may frustrate the public and lead to more executions or, when the execution rates remains constant, undermine the support of the death penalty.

Whether or not there is a general deterrent effect, there is certainly a diminishing effect on crime "since execution eliminates categorically the possibility of recidivism" (Ehrlich, 1975a).

2.2.3 Drugs

> *"Would you tell me, please, which way I ought to go from here?" "That depends a good deal on where you want to get to", said the Cat. "I don't much care where...", said Alice. "Then it doesn't matter which way you go", said the Cat. "So long as I get somewhere", Alice added as an explanation. "Oh, you're sure to do that", said the Cat, "If you only walk long enough."*
>
> *Lewis Caroll, Alice in Wonderland, 1865*

Drug consumption can lead to increasing crime rates. Either due to physical or psychic consequences of the drug consumption itself or crimes committed to obtain the required money to buy drugs. Additionally, solving problems between drug dealers or consumers cannot be moderated by the police and may lead to more violent crime (Resignato, 2000). Using a German state panel, Entorf and Winker (2002) find that drug crimes significantly influence rape and property crimes. As a consequence, ignoring the drug issue in a study of deterrence may introduce a small positive bias. However, Resignato (2000) cannot find such effects in a study of 24 U.S. cities.

Since institutional resources are limited, redeploying police resources from other offenses towards drug related crimes may reduce the latter but increase all other crimes (Benson et al., 2001; Resignato, 2000). Moreover, any distortion of the drug market equilibrium is assumed to be accompanied by violent crime. Rasmussen et al. (1993) and Benson et al. (1998) also study this effect by analyzing the interdependencies between violent, Index I or Index II crimes and drug related arrests in Florida. Indeed, they conclude that the increased efforts on the war against drugs lead to an increase of other offenses. Benson et al. (1992) argue in the same fashion that only a small portion of drug-criminals commit property crimes (20%) and react to changes of deterrence measures in regard to property crimes. Since in the USA arrestees of drug crimes are generally imprisoned, overcrowded prisons might lower deterrence overall and thus lead to more property crimes. This could even imply more illegal income for drug users and more drug consumption. The same is observed for the homicide rate. An increase of drug enforcement activity by 1% is associated with an increase of the homicide rate by about $0.1\% - 0.17\%$ (Brumm and Cloninger, 1995a). Therefore, it makes more sense to concentrate on property crimes which would decrease the income and consumption of at least those drug users who rely on property crime.

On the other hand, Kaplan (1983) argues that it may be just the illegality of drugs which leads to more illegal activities and not the drug usage itself. Illegality implies more problems getting or remain employed, forces prices upwards and drug users into a criminal subculture. Furthermore, recidivism is much lower for those convicted for drug offenses only than those who also engage in non-drug crimes (Kim et al., 1993). However, measuring drug consumption might is difficult (Entorf and Winker, 2002). Using drug casualties is unappropriate (because it depends on thy quality and type of drugs, lags, etc.) as well as is the number of new drug consumers (inaccurate approximation). They resort to the official number of drug offenses although this is also not perfect.

Another aspect is that the police may succumb to other incentives. In the USA the Federal Comprehensive Crime Act of 1984 allowed the police to keep the proceeds from assets seized. Mast et al. (2000) find that the police spend, as a consequence, too much resources on drug enforcement activities instead of drug treatment. In this context it should be mentioned that according to Swimmer (1974a), too much money is spent on the police in cities with low and medium levels of crime while police expenditures are too low in high-crime cities.

Silverman and Spruill (1977) point out that the price of cocaine explains much of the variation in crime. The type of drug and its consumer is also important when considering deterrence. DeSimone and Farrelly (2003) analyze the effectiveness of increases in the price and enforcement of marijuana and cocaine on the consumption by male adults and youths. They find that adults - juvenile drug demand is not price sensitive at all - decrease consumption of both drugs when the price of cocaine increases (see also DeSimone (2001)); increasing prices for marijuana has no effect at all. Enforcement, measured by the probability of arrest for possession of these drugs, seem to be effective. Cross-arrest effects indicate that these drugs are complementary goods. Farrelly et al. (1999) conclude that even marijuana and legal drugs are complementary goods and that deterrence is affective for adults but not for youths. This is supported by McGeorge and Aitken (1997) who study Australian students in states where the usage of cannabis is legal and states where it is illegal. No differences were found at all - only the knowledge of related laws and penalties were somewhat inaccurate. Lenton (2000) summarizes some studies about the criminalization of possessing and using marijuana in Australia and finds that there is neither a general nor a specific deterrent effect. Therefore, the famous *false-signal effect* of the legalization of marijuana is invalid (because nobody reacts to this signal) and the stigma of such a drug record has only negative effects on the further life of the delinquent. Burkett and Jensen (1975) point out that, surveying senior-class students from three U.S. high schools, conventional ties and peer involvement may be even more important than the subjective probability of detection. Nevertheless, that probability is negatively influencing the consumption frequency significantly.

A very good testing ground can be found in the U.S. military and their no-tolerance drug policy. The personnel is randomly tested and any drug detection is followed by immediate dismissal. These constellation should provide very clear results. However, while the harsh punishment significantly deters (Mehay and Pacula, 1999), the effects are not as strong as expected. In a very similar setting, Borack (1998) compares the U.S. Navy and the general population and concludes that the random testing in the Navy deters much (over 50%) of the potential drug usage.

2.2.4 Tax Evasion

> *In fact, the puzzle of tax compliance is that most people continue to pay their taxes. [. . .] Although it is clear that detection and punishment affect compliance to a degree, it is equally clear that these factors cannot explain all, or even most, tax compliance behavior.*
>
> *Alm et al. (1992b)*

In general, it is assumed that tax evasion has a fairly low detection probability which is usually overestimated by the public and is punishable by fines or imprisonment. As for other offenses, the results for tax evasion are mixed. Feld and Frey (2004) remark that "taxpayers should evade more than they actually do".

As is reported now and then for usual crimes, Isachsen et al. (1985) find, using Norwegian data, deterrent effects of the detection probability and the severity of punishment. However, Sheffrin and Triest (1992) point out that the effect of the detection probability may be overestimated when the interaction between authorities and tax payers is not considered. They survey U.S. tax payers and find that quarrels with the authority slightly deters but also reduces the subjective probability of future detection. Subsequently this might lower future tax returns and reduce any stigma effect. Additionally, menacing actions by the authority may backfire while factual and technical arguments foster tax honesty. This complies with the notion by Anderson and Stafford (2003) that people tend to to follow the "lightning doesn't strike twice" saying: once an offense is detected the perceived probability of a future detection is reduced by a large margin. The availability thesis (Spicer and Hero, 1985) implies a contrary conclusion: people adjust their subjective probability of detection upwards when they are caught and downwards otherwise.

Alm et al. (1990b) disentangle regular tax payments (including tax avoidance) and tax evasion in Jamaica in regard to tax rates and deterrence. They find that increased deterrence leads to a reallocation from tax evasion to tax avoidance accompanied with a small reduction in the overall tax revenues. Another aspect is that increasing taxes may even increase all (property) crimes because taxes are only paid for legal activities which increases the incentives for illegal income generation (Virén, 2001).

It may also be important to distinguish more and less severe tax evasion. Smith (1992) finds that the probability to be caught for minor tax evasion does deter while that for severe tax evasion does not. Deterrence seems to work very well when the tax payers have accurate information about the detection probability but does not work at all when this information is not available (Spicer and Thomas, 1982). Thurman (1991) distinguishes between the "sin of omission" and "sin of commission": while the first relates to underreporting (omitting some income) and may be passive, the latter refers to the active action of overstating deduction. All in all, he can't find any deterrent effects but overstated deduction is affected by the perceived threat of guilt. Sheffrin and Triest (1992) conclude that self-reports are more suited for research than audit-reports because many violations happen by mistake. Likewise, Beron et al. (1988) classify overstated deductions as easy and underreporting as hard to detect. Moreover, they give a short description of how the Internal Revenue Service (IRS) detects possible tax evaders by looking for outliers in the reporting behavior. They find a significant but minor deterrent effect of the audit probability. Bosco and Mittone (1997) made an experiment with 60 Italian students and find no effect of the audit probability. Furthermore, the amount of evaded taxes is too high when the audit probability is of medium size and the main influencing factors are moral necessities. Pudney et al. (2000b) distinguish between the decision to evade and the evaded amount. Using an experiment with 270 Turkish people they conclude that the tax rate influences both the probability and the evaded amount while the expected sanctions only affect the former. Similar evidence is found by Benjamini and Maital (1985) who find that increasing the tax rate fosters tax evasion while the probability of detection has no major impact. They add that the probability of detection is underestimated and that risk averse people evade less. Bosco and Mittone (1997) remark that the degree of risk aversion is more important than the audit probability.

However, Hessing et al. (1988) cast some doubt on results based on surveys. They use both a survey and official data from the Netherlands and come to the conclusion that there is no correlation between them and their implications contradict each other. A different approach is to measure tax evasion by comparing the national accounts measure of primary income and the income reported to the tax authorities (Feld and Frey, 2004).

Tax evasion or the attitude towards free riding[17] is often studied with laboratory experiments (usually playing games with students). Evading taxes does not exclude anyone from the consumption of public goods. As an example, Anderson and Stafford (2003) let U.S. students spend money on a public good or keep it. The public good had a doubled return, distributed among all students. Keeping was detected and punished at different rates and with various sentences (the parameters of the experiment). In theory, depending on the expected punishment, free riding or keep nothing are dominant strategies. They find that the probability and severity of punishment are effective but not the expected punishment. They conclude that probability and severity are not substitutes and these results are applicable to tax evasion. Indeed, Friedland et al. (1978) also report in an experiment - keeping the expected punishment constant - that monetary fines deter while the probability of detection does not. Park and Hyun (2003) use an experiment with 15 students in South Korea and find that the penalty rate deters more than the audit rate. Furthermore, tax education improves tax honesty and there is an obvious tendency towards free riding. Small detection probabilities are overestimated and, even when these are zero, some participants remain honest (Alm et al., 1992b). When players are allowed to punish other players for free riding, the free-riding strategy remains

[17] "One would like everyone else to pay true tax while evading oneself." (Benjamini and Maital, 1985)

dominant. Fehr and Gächter (2000) made an experiment with Swiss students and find that these do - contrary to theory - make fairly often use of the punishment option, thereby reaching even more efficient states. The cooperation is even higher when the players know each other. However, Bosco and Mittone (1997) point out that evading taxes may even increase welfare when the taxpayer is convinced that he is paying too much taxes relative to the provided public goods.

Benjamini and Maital (1985) emphasize the phenomenon of sub-certainty in the context of tax evasion: the sum of the weighted probabilities $V(p)$ and $V(1-p)$ is smaller than one. Furthermore, very low probabilities are often overweighed. A similar problem is reported by Casey and Scholz (1991) who observe that people do weigh single probabilities equally to the product of multiple probabilities (of similar size). In an experiment the subjects were confronted with either a single punishment probability or several probabilities (detection-, conviction- and punishment probability). While both probabilities of punishment were equal the latter was perceived to be higher. Spicer and Hero (1985) note that - under great uncertainty - people tend to replace rational optimization by some rule of thumb.

Analyzing experiments with U.S. students Alm et al. (1990a) find that tax amnesties as well as the penalty for tax evasion has not the intended effects. Neither does an increased penalty nor does an amnesty foster tax revenues. However, a combined strategy (a tax amnesty followed by increased penalties) does lead to significantly increased tax revenues. According to Ritsema et al. (2003), people directly affected by an amnesty can be categorized using three groups (intentional, unintentional and neutral) who behave differently. While the lack of money is a main reason of tax evasion for intentional noncompliance, perceived unfairness by the system is important for the neutral group. No such main influences are found for unintentional noncompliance.

When dealing with tax evasion it is also important to consider the general attitude towards the tax system. Fairness (the treatment by the authorities) and reciprocity (the individual acts as the masses do) influence tax evasion behavior. A *respectful* treatment by the authorities can bolster tax morale while an *authoritarian* treatment can crowd it out (Feld and Frey, 2004). While fairness has a positive influence it mainly affects those people who had already contact with the authorities (Smith, 1992) but has only little direct effects on all others (Scott and Grasmick, 1981). Normative responsibility and deterrence are found to be complementary factors (Smith, 1992). However, Alm et al. (1992b) find no difference in the formal addressing, with and without explicit reference to tax evasion, indicating the absence of a normative influence. In a small experiment with eight students Alm et al. (1992b) also analyze the formulation of the experiment (one neutral and one explicitly mentioning tax evasion) but found no differences and thus neglect moral factors. Perceived fairness may also be negatively affected by the probability of detection and thus even lower the tax revenue (Frey and Feld, 2002). Fortin et al. (2004) employ an experiment with Canadian students and - avoiding unobserved heterogeneity and identification problems - come to the conclusion that the audit probability does not deter and only fairness effects are significant. The general attitude towards the tax system is also studied by Steenbergen et al. (1992). Although the knowledge of a tax reform and its social consequences are found to be important, the overall effects are negligible.

While, in general, richer people are found to evade more, people from the middle class do almost not evade at all (Beron et al., 1988). However, Alm et al. (1992a) find in an experiment with U.S. students that tax evasion increases when income decreases or tax rates increase. Spicer and Lundstedt (1976) point out that some tax evasion may occur to balance perceived unfairness because many believe that only the rich and companies are responsible for tax deficits. Attitudes and behavior in regard to tax evasion may also vary between different groups. Gërxhani and Schram (2006) compare pupils, students, teacher and university staff in the Netherlands and Albania. They find that the audit probability does deter in the Netherlands but not in Albania. Furthermore, the differences within the groups are in some cases larger than across the two countries.

2.2.5 Environmental Offenses

They are ill discoverers that think there is no land, when they can see nothing but sea.

Francis Bacon, The Advancement of Learning, 1605

The most frequently studied environmental offense is the spillage of oil in rivers, harbors or oceans. Weber and Crew (2000) study oil spillage in the ocean. They use the fines and the probability and celerity of punishment as deterrents and find that all but the probability are effective. Anderson and Talley (1995) distinguish the detection probability by the vessel type and conclude that it is only effective for barges but not for tankers while shipping under the U.S. flag is the most important determinant. Epple and Visscher (1984) employ data of the U.S. coastguard on an individual and aggregated level. Using the spill size and the spill rate they find strong enforcement effects. Cohen (1987) also studies tankers and barges and the effects of the surveillance of oil transfers, patrols and inspections. In regard to the size of the spills the former are effective deterrents while inspections are not. Gawande and Wheeler (1999) use the U.S. coast guard as a representative for a non-profit governmental organization. They use poisson regressions of accidents (fatal, non-fatal and pollution) on inspection-hours and conclude that inspections deter and do not depend on the specific type of inspection (hull, machine or books). Grau and Groves (1997) emphasize that the detection probability is the most important deterrent and that fines play no role because they are too low. Weber and Crew (2000) report exactly the opposite: while fines and the celerity of punishment effectively deter, the probability of detection does not.

Aside from oil spillage, most other studies are concerned with waste regulations, e.g., Stafford (2002) or Magat and Viscusi (1990), or water pollution (Storey, 1979). Magat and Viscusi (1990) find strong immediate deterrent effects of inspections in the pulp and paper industry on pollution levels and the rate of compliance. Storey (1979) finds only very mixed results for water pollution in England and Wales (where the “consent on local level” approach is used in contrast to the U.S. approach of “one law for all”) with monetary fines seeming to be the most effective deterrent. It may be argued that some deterrence - especially for companies with a good reputation - result from reputation effects. However, according to Karpoff et al. (1998) this seems not to be the case; at least not for listed companies. They find that legal penalties, which do not correlate with firm size and are hard to predict, are decisive and that these legal penalties reduce the share values while the reputation effect is negligible.

There are even some studies dealing with rather odd topics like the attitude towards fishing quotas (Hatcher et al., 1999). They use a survey of English fishermen and find that - considering the feedback effect between the subjective probability of detection and exceeding the quota - the probability is a significant deterrent. Moral and peer influences are also very important while the attitude towards the legitimacy of the quotas is not. Similar results are found by Furlong (1991) who surveys Canadian fishermen. The subjective detection probability and the conviction rate deter; although not significant, the expected punishment is more effective than a licence revocation.

2.2.6 Drunk Driving

In nature there are no rewards or punishments; there are consequences.

Horace A. Vachell, The Face of Clay, 1906

Although economists are normally less skeptical towards the deterrence doctrine than sociologists and criminologists, this relationship is reversed for DUI offenses (Benson et al., 2000). This may result from the notion that the perceived probabilities are so low in practice that the reactions should be inelastic. Aside from the large number of potentially drunk drivers there might be another reason for low detection and arrest probabilities in the USA: most people who drink and drive come from “middle income groups with more political ties”. Policemen might be more reluctant to arrest them, especially for states without breath-test laws which allow immediate legal tests of the Blood Alcohol Content (BAC), as Saffer and Chaloupka (1989) point out.

Nevertheless, there are basically three classes of instruments for the government to influence drunk driving (Stout et al., 2000): administrative regulation, criminal laws and tort liability. Administrative regulation can affect the behavior of consumers via regulation of the alcohol industry (which influence the consumption of alcohol). This can be achieved by a monopoly control system (state owned stores or licensing), advertising practises or by influencing the price of alcohol (via taxes). Criminal laws are meant to deter, although some also incapacitate (either by imprisoning or license revocation). Among such laws are mandatory minimum punishments (fines or jail or license revocation), fines, ban of open containers of alcohol in vehicles, administrative per se laws, sobriety checkpoints and many others. These are meant to increase the severity, certainty or celerity of punishment, thereby deterring DUI. Tort liability laws impose “civil penalties, usually in the form of monetary damages, on those who are found to be at fault in causing harm”. For example, dram shop and social host laws allow people injured by an alcohol-impaired person to sue those who have served the alcohol to that person.

BAC limits are widely used to define and reduce drunken driving. Mann et al. (2001) give an overview of studies analyzing the introduction or change of BAC limits in an international context. While there is a great variety in results, there seems to be some beneficial effect in almost all cases. Although many of these effects are only small or temporary, there are also some lasting effects to be found. Nonetheless, these effects seem to be attributable to general deterrence and are not only applicable to driver with BAC levels around the legal limit. Stout et al. (2000) cite several studies which report mixed results for administrative regulation and criminal laws. However, they are confident that tort liability is a useful tool to reduce DUI.

Similar to the cascade from detection to punishment (refer to page 43), the same principle should apply to laws in the context of DUI-offenses. As Saffer and Chaloupka (1989) note, laws which increase the detection probability (e.g., breath-test laws, reduced BAC limits, sobriety checkpoints, etc.) should be, in theory, more effective than subsequent laws (e.g., minimum terms, licence revocation, etc.).

Most studies about DUI can be classified into two categories: surveys or natural experiments. The latter usually consist of an analysis of one or more laws introduced at a certain point in time and its influence on accidents or DUI-related arrests (usually with dummies in a time series analysis). For example, Evans et al. (1991); Mann et al. (2002); Ross and Klette (1995); Chaloupka et al. (1993); Rogers and Schoenig (1994); Foss et al. (1998) and Glass (1968) use such regime changes. However, as Maghsoodloo et al. (1988) point out, such new laws are likely to be effective only when the perceived difference (before and after the law change) is large enough and an education effect can take place. Surveys are often used in conjunction with young people who are generally assumed to be more prone to drunken driving. For example, Rabow et al. (1987) ask young college students about their drinking and driving habits. Classic deterrence variables like the probability of detection or knowing DUI-victims are not significant. However, beside social influences, the knowledge of reducing the risk of apprehension (e.g., knowing how to cover ones own drunkenness) is effective. Richardson (2003) considers potential influences of the subjective probability of detection, pointing out that drunken people are, per definition, not fully rational and are no longer capable of assessing the risks, costs and probability of being caught. Among several new laws introduced at state level, only the introduction of sobriety checkpoints found to be significant. Licence revocation is considered more severe than fines or imprisonment. Nevertheless, none of these punishments are significantly influencing the individual drinking-and-driving habits.

Many authors argue that laws and other deterrents only apply to a small group of potential drunk drivers (Houston and Richardson, 2004; Soper and Thompson, 1990). While non-drinkers do not react per definition, hardcore drinkers are also assumed to be unaffected in their habits (Yu, 2000; Foss et al., 1998). Thus, only those people who belong to neither of these two groups can be deterred (Mann et al., 2003). Indeed, studying U.S. high school seniors who had drunk in the last 30 days, Grosvenor et al. (1999) find that only those react to the probability of detection (but not to the severity of punishment) who are characterized as binge drinkers while all others do not react at all. Similarly, in the case of general traffic laws, Ross et al. (1990) argue that such laws mostly affect people who already drive carefully and not the targeted group which is rather immune to such threats. Furthermore, the knowledge of DUI-laws is, especially in the USA, quite low and heterogeneously distributed. According to Berger and Marelich (1997), only 30% know the actual BAC limit while more than 80% of all Australians and Norwegians do. Although the sanctions in the USA are more severe than in Norway these are judged less relevant in the USA. Although such knowledge increased in the eighties, this was accompanied by a large reduction in the perceived probability of detection (Snortum and Berger, 1989). The knowledge of DUI-laws is extraordinarily poor for non-drinkers and notorious drinkers (Kenkel and Koch, 2001). Foss et al. (1998) report that heavy drinkers may recognize law changes and their effect on the probability and severity of punishment but think that they are not affected. However, Berger and Snortum (1986) do not find any correlation between this knowledge and drunken driving. A representative survey of Canadian people analyzed by Wilson and Jonah (1985) indicates that the alcohol consumption of the past seven days is the best predictor for DUI, as well as that non-drinking drivers have the highest and impaired drinking driver have the lowest risk perception. They conclude that "impaired driving may be just one [...] syndrome typified by high-risk behavior". Similarly, Mann et al. (1996) conclude that alcohol consumption is the most important factor affecting fatal accidents while membership of Alcoholics Anonymous is negatively related to the latter.

Others argue that positive general deterrence may be more important than negative general deterrence. In the USA grassroots organizations like Mothers Against Drunk Driving (MADD, founded in 1980) have increased public awareness of the DUI-problem and initiated an anticipatory deterrent effect (Rogers and Schoenig, 1994). According to Snortum and Berger (1989), the government began to take concrete action in 1983. Many laws to diminish the DUI-problem followed and, depending on the employed data set, distinguishing these two effects can be difficult. Evans et al. (1991) find weak evidence for some combinations of new laws but conclude that the reduction of accidents was more likely to be caused by the increased awareness of the general public. Berger and Snortum (1986) claim that the moral attitude is more important than deterrence. Berger and Marelich (1997) also emphasize that it is important to differentiate between general prevention in the context of deterrence and norms (for example, Norwegians disapprove of drunken driving in general while Americans do not). Snortum and Berger (1989) add that men in the USA "have traditionally carried both a *social obligation* to drink and a *social responsibility* to drive". He also concludes that laws seem to have only short term effects (the probability of punishment is overestimated in the beginning and then wears off) while there is a long term educational effect. While summing up studies dealing with jail terms and drunken driving (in the USA) in regard to specific and general deterrence, Voas (1986) notes that "drinking and driving was socially acceptable and juries tended to be lax in their treatment of offenders". He also remarks that the "more subtle role of the jail sentence in conditioning public attitudes toward drinking and driving by raising the penalty for committing this offense" had not been studied at all.

Vingilis et al. (1988) use impulse functions and monthly data on fatal accidents with and without alcohol. Fatal accidents are well suited for this purpose because everyone involved has to be tested for alcohol. They study the effect of a law in Ontario which requires a drunken driver to be deprived of his driving licence for 12 hours. They conclude that there is probably a deterrent effect but it quickly diminishes after a few months. The rapid deterioration of a deterrent effect is a result which is also found in many other studies. Mann et al. (2002) also use Canadian data (from Ontario) where drunken drivers are deprived of their licence for 90 days. Although the results are somewhat mixed they tend to favor a deterrent effect. Fatal accidents (total, at night, youths and youths at night) are also studied by Saffer and Chaloupka (1989) for U.S. states. Using cross section and fixed effects estimations they conclude that all analyzed breath testing laws are significantly deterring drunk driving. These results are contradicted by Ruhm (1996) who finds no such effects but argues that an omitted variable bias (omitted influence of grass root organizations like MADD and beer taxes) are responsible for an overestimation of the deterrent effect. The most important variable he - as well as Chaloupka et al. (1993) - identifies are beer taxes. More expensive beer leads to less drinking which implies less drinking and driving. Berger and Snortum (1986) also state that the general consumption of alcohol is more important than deterrence and, according to Ruhm (1996), more robust to alternative specifications. Using a large bunch of deterrence variables in an U.S. panel Whetten-Goldstein et al. (2000) also find no obvious effects on fatal accidents (except administrative per se laws for adults and minimum fines for youths). Similarly, Mullahy and Sindelar (1994) use individual and state data from the USA and conclude that many law adjustments had no effect (not even short termed). Only fines, licence revocation and beer taxes have large deterrent effects. However, using data from the NCHS (in regard to alcohol related mortality), Sloan et al. (1994) find no effect of fines and licence revocation (and minimum jail terms as well) while they find a deterrent effect of dram shop laws and the police on fatal accidents.

Levitt and Porter (2001) present a model which is capable of deriving the participation rate in fatal accidents - without knowing the actual numbers - under the assumption of equal mixing: homogeneity in time (year, weekend, hour) and regions. With that model the increased risk of a drunken driver of being involved in a fatal accident can be estimated. Taxes on alcohol and fines are studied and they report that punishment is effective for those individuals without prior records. While those already convicted are more careful and less noticeable. Another unusual approach is taken by Sloan and Githens (1994) who analyze premium penalties imposed by insurers for drunk driving and other chargeable accidents. Using a survey of insurers they find that “imposing premium surcharges for a charge of drunk driving has a significant deterrent effect on the probability of drinking and driving”.

Another aspect is the implementation of new or harsher laws in practice. For example, Ross (1987a) reports that the introduction of an administrative licence revocation law (to increase the swiftness of punishment) had a small deterrent effect which could have been greater if the police and judges would have made more use of it. A large portion of his study deals with the acceptance and usage of that law by the police and judges and the - at best, mediocre - coverage by the media. Soper and Thompson (1990) also report that deterrence is only found to be effective when it is implemented efficiently; e.g., by training of the police accompanied with a good information policy and coordination. Harsher sentences and fines may even have some perverse effect because offender will more often insist on trials and appeals and thus lengthen the time frame between delinquency and punishment (Ross et al., 1990), thereby reducing the deterrent effect of the celerity of punishment.

Although U.S. data is used very often, there are several studies analyzing other countries. Especially, Scandinavian countries are quite often subject of DUI-studies because “it is common knowledge that among the advanced industrial nations, they have some of the most restrictive regulations in regard to the availability of alcoholic beverages for sale, and that they maintain some of the most punitive and rigorously enforced laws with respect to drunkenness and driving” (Votey, 1978). Ross and Klette (1995) use the “abandonment of mandatory jail for impaired drivers in Norway and Sweden” but do not find any significant effect with an interrupted time series approach. Ross (1975) uses time series from Norway and Sweden and, finding no effects, dismisses the notion that the DUI laws in Scandinavia have a measurable impact as the “the Scandinavian Myth” implies. However, Votey and Shapiro (1983) use Swedish data on a monthly basis and find that the expected jail time and fines have some deterrent effect while the arrest rate and licence revocation have large deterrent effects. Using data from Norway and Sweden, Votey (1978) finds deterrent effects of the probability of punishment as well.

The costs of combating DUI are addressed in Kenkel (1993). Using survey and panel data he finds that indirect (for example, taxes on alcohol) and direct measures (e.g., minimum sentence, checkpoints) are effective but costly. Especially, laws and taxes usually affect all people - even those who never drink and drive. On the other hand, some deterrent measures (like fines) work by generating costs (almost) only for the offender. Likewise, Votey and Shapiro (1983) point out that licence revocation lays the burden on the individual offender while jail time, for example, has to be paid by the whole society. Levitt and Porter (2001) calculate the costs of accidents (under the influence of alcohol, using U.S. data) to be 16-30 cents per mile. If these external cost were internalized, each arrest would have to be accompanied by an 8000$ fine.

2.2.7 Crime Switching

To see what is general in what is particular, and what is permanent in what is transitory, is the aim of scientific thought.
Alfred N. Whitehead, An Introduction to Mathematics, 1911

Koskela and Virén (1997) present a crime-switching model in which criminals switch from on type of crime to another, if the expected punishment makes one type offense less attractive. They test their model with Finnish crime data (vehicle theft and robberies) and relative arrest and punishment rates. They find deterrent effects which support their hypothesis. In another study, Koskela and Virén (1994) partition the general populations into criminals and non-criminals by their productivity in the legal sector (using wages, tax rates, social payments, punishment, etc.). Using Finnish data of auto thefts, they find “that the rate of return from illegal activity has a positive, and the apprehension rate and the severity of punishment a negative, effect on auto thefts both in the short run and in the long run". Again, they find evidence for the applicability of their crime-switching model. This view is supported by Holtmann and Yap (1978) who use crime and imprisonment rates in the USA for robbery, burglary and theft. Although not the focus of the study, Fabrikant (1979) also finds such changes in a mixture of offenses when the probability of failure (of an offense) changes. In a similar way, Levitt (1998c) argues that violent crimes are substitutes for each other (as well as property crimes) and uses this assumption to distinguish deterrence from incapacitation. In contrast, no substitution effect is found by Merrifield (1997), if the expected return of the former offense diminishes.

2.2.8 Youths

Not everything that can be counted counts, and not everything that counts can be counted.

Albert Einstein

Generally youths, in contrast to adults, are threatened with less severe punishments. While these form of clemency is accepted world wide, there may be unintended consequences. Levitt (1998a) points out that juvenile violent crime did grow twice as quickly as that of adults in the eighties and nineties ins the USA. He finds that sixty percent of this difference can be explained by the more lenient punishment of juvenile offenders. There is a sharp drop when youths reach adulthood but there appears to be no strong relationship between the criminal involvement before and after reaching adulthood. In Levitt and Lochner (2001) they employ a different set of data and find, using a differences-in-differences model, that crime is reduced in the transition phase between the age of 17 and 18. However, these reductions are not significant in most cases. In regard to property crime, there is weak evidence that youth- and adult crimes are substitutes (youths replaces adults) while there is also weak evidence that they are complements (adults are role models for youths) in regard to violent crime (Levitt, 1998a).

Mocan and Rees (2005) cite several studies which note that youths in the USA do not (or only very weakly) react to deterrence measures and become more violent in their criminal actions. They use representative micro data of 1995 and come to the conclusion that "juveniles do respond to incentives and sanctions as predicted by economic theory". Furthermore, the ratio between male and female juvenile offenders is about two to one.

2.2.9 Other Offenses and Laws

The plural of anecdote is not data.

Roger Brinner

Besides the offenses already mentioned, several other very interesting types are considered in the literature.

Gun Laws

Canada introduced a law in 1977 to restrict the carrying of firearms. In principal, such laws can have two different consequences: potential offenders may have less access to firearms which would lead to fewer armed robberies; or more victims are unarmed and become easier to rob which would lead to an increase of armed robberies. Mauser and Maki (2003) study robbery, armed robbery and armed robbery with firearms in Canada during that time. They use all combinations of covariates and do not find, using OLS, any consistent results (halve of the results support deterrence theory, the other halve does not). GLS showed a slight positive effect of the law, implying that robbers, who tend to ignore such laws, have larger chances to encounter unarmed victims. The probability and severity of punishment are correctly signed but not significant.

John Lott and David Mustard are probably the most popular advocates for "right-to-carry concealed handgun gun" laws in the scientific literature[18]. In Lott and Mustard (1997) they use county panel data of the USA and find, relying on dummy variables, large significant effects of such laws for Index I crimes and substitution effects into property crimes. Additionally, they conclude that the monetary gains (by deterred crimes) are much larger than the marginal monetary drawbacks (e.g., deadly incidents).

Dezhbakhsh and Rubin (1998) reevaluate Lott's data and refrain from using dummy variables to analyze the effect of such laws. Instead, they compare the coefficients of the exogenous variables of two different regressions: counties with and without such laws. If the change of the parameters is significant, concealed-handgun-laws have an effect. Overall, only murder is slightly reduced by the usage of such gun laws for some states (which do not have such a law). The positive effect on many robbery rates is explained by the low potential threat of concealed guns because robbers are often already armed and thus protected against armed resistance. In another study, Dezhbakhsh et al. (2001) use Lott's data and find that NRA (National Rifle Association) memberships are associated with higher murder rates. Cook and Ludwig (2002) add that guns provide a valuable loot for offenders and, using data from the UCR and NCVS, conclude that the deterrent effect is outweighed by the increased incentives. A different approach is implemented by Ludwig (1998). He utilizes the minimum age required to possess a gun and the age information of the cleared homicides. If these gun laws do deter, the homicide rates for adults should decrease while the rates for juveniles should be unaffected (or increase in the case of replacement). Using a differences-in-differences-in-differences approach, he cannot find any deterrent effect. Summarizing the literature, Levitt (2004) concentrates on the receding crime in the USA in the 90s and concludes that gun control laws and carrying concealed guns laws had no effect.

In an a answer to Black and Nagin (1998), who point out several flaws in Lott's work, Lott harshly rejects their criticism and gives a reevaluation of his data set and shows that violent crime increased until the gun laws were passed and than sharply dropped after a short lag (Lott, 1998). Very strong deterrent effects are also found by Bronars and Lott (1998) who add that there also

[18] A large portion of the literature about this topic is somehow related to Lott.

exist serious spillover effects because "criminals tend to move across communities more readily in response to changes in concealed handgun laws than in response to changes in arrest rates". Studying safe-storage gun laws Lott and Whitley (2001) find no evidence that such laws "reduce either juvenile accidental gun deaths or suicides". Contrarily, they report that these laws "impair people's ability to use guns defensively" and even increase violent and property crime.

Sports

Several authors utilize changes of rules in the world of sports. McCormick and Tollison (1984) use an annual basketball tournament in the United States (1954-1983) to draw conclusions from sports concerning the effect of the police on the arrest rate and subsequently the crime rate. They analyze the introduction of a third referee in 1979 and find that the number of fouls are reduced by 34% while the number of false decisions is also reduced. Additionally, the variance of the results diminish and the games end with higher scores. They conclude that more police should indeed reduce crime.

Studying the introduction of a second referee in the National Hockey League (NHL) in some games of the season 1998/99, Allen (2002) reports that the number of detected fouls increased. This effect is declared to be a reporting effect (more "offenses" can be detected with a second referee). The model used can explain non-violent fouls very well while violent fouls seem to happen randomly. The failure to detect any deterrent effect may result from the usage of inadequate data as Levitt (2002a) points out. In his model, using the same season, the number of actual minor fouls did not change significantly while, at the same time, the second referee did not change the probability of detecting such fouls as well. In the season 1999/2000, the second referee was used in randomly selected games. Using an instrumentation (the actual number of games played with one referee is used as an instrument because the number was equal for all teams at the end of the season) approach, Heckelman and Yates (2003) distinguish monitor and deterrent effects. While they find strong monitoring effects for fouls, penalty minutes, minor and major penalties, no deterrent effects are found.

The English Premiere League is used by Witt (2005) to identify deterrent effects of the tightening of rules in 1998. He finds that the number of red cards remained constant while the number of yellow cards significantly increased. Thus, he argues that there was a deterrent effect because increasing the pool of fouls punishable by a red card (e.g., tackle from behind) lead to a substitution by less severely punishable offenses.

Corruption

Since reliable data about corruption is difficult to obtain, there are only some isolated studies. Vinod (1999) uses the corruption perception index (based on a survey of "business people, risk analysts, and the public"). He employs the ratio of judges to population as the independent variable and concludes that increasing the efficiency of the legal system indeed decreases corruption.

Goel and Rich (1989) use the ratio of officials convicted for bribery to all officials. Both conditional conviction rate and the sentence length are significant while police expenditures are not. Furthermore, bribery increases when unemployment or consumption increase or wages decrease. In Goel and Nelson (1998) they use the conviction rate as the dependent variable and find “that government size, in particular spending by state governments, does indeed have a strong positive influence on corruption” while only the police expenditures and the number of other justice department employees are significant deterrents.

Cheating in the Classroom

Although “cheating in the classroom” is certainly a very specialized topic, there are several studies which deal with this type of malpractice. Results are ambiguous. Mixon and Mixon (1996) asked economic students and find mixed results for the probability of punishment and correctly (i.e., supporting the deterrence hypothesis) signed variables for the severity of punishment. No deterrent effects are found by Bunn et al. (1992) using another survey of economic students. Nevertheless, other results are that peer influences (knowing of other students who cheat) and bad grades (penalty is less threatening) foster cheating behavior.

By contrast, Houston (1983) report, asking psychology students, that good students cheat more but do react to threats of punishment while bad students cheat less but do not react to threats. However, the threats are only effective when their severity passes some threshold. The gender os the students is not relevant.

Other Offenses

While most studies deal with standard offenses or deterrence measures, there are some studies which are out of the ordinary. DiPasquale and Glaeser (1998) analyze riots in U.S. cities in the sixties. As expected, ethnic reasons are important factors, not police variables. Sirakaya and Uysal (1997) analyze the compliance of tourists with local guidelines in the USA, Canada and Ecuador. Using three factors - deterrence, education and awards/motivation - they find that only education is relevant. Fujii and Mak (1979) and Fujii and Mak (1980) use data from Hawaii and find that increasing tourism also increases crime (especially burglary and rape) in general and that crime is more concentrated in areas populated by tourists. They advise to reduce the overall number of tourists and to concentrate more on wealthier visitors. Tourists in Hawaii are also studied by Ghali et al. (1983) who come to similar conclusions but find such effects for theft but not for burglary.

Although arson is classified as an Index I crime, almost no study considers it. Cloninger (1981) and Cloninger (1990) are exceptions and he concludes that arson, interpreted as an instrument used for insurance fraud, is significantly deterred by the corresponding clearance rate. Another unusual study is that of Braithwaite and Makkai (1991). They analyze corporate deterrence by surveying managers of Australian nursing homes and do not find any influence of the severity of punishment but some effects for the probability of punishment on official compliance. Landes

(1978) studies airplane hijacking in the USA. The arrest rate and sentence length are found to be significant deterring hijacking while the conditional conviction rate is not. He distinguishes between ex ante (e.g., screenings) and ex post (e.g., marshals) deterrence measures and considers a trend using international data.

Insurances can have the perverse effect of increasing the number of insured events (e.g., unemployment insurance may increase unemployment). However, Cameron (1989) uses state-data and concludes that no such effects can be found for victim compensation rewards for rape and aggravated assault. Clarke (1966) studies school boy absconding and its corporal punishment in 1960-1964 in Kingswood, United Kingdom. During that time absconding school was punished by strokes with a cane. He finds that juniors are not affected while seniors are. This is explained by psychic pressure (homesickness, indisposition) which is more dominant for younger boys.

Commonly, offenders are assumed to be aware of their illegal or undesirable behavior. This may not be true in all cases, even if they are fully aware of the consequences because the punishment may have refined their point of view about the offense. Gneezy and Rustichini (2000) report an experiment they made in an Israeli kindergarten. Many parents picked up their children too late so that the kindergarten teachers had to stay longer. Therefore, a fine was introduced for parents who showed up too late. Instead of relieving the problem, the occurrences of late parents even increased because they interpreted the fine as a price which they payed for coming too late. Since the fines were moderate, they seemed to be smaller than the opportunity costs of arriving earlier. They conclude that this reasoning might be transferred to similar problems like tax evasion. It seems reasonable that other misdemeanors, like false parking or speeding, may also apply to this scenario.

2.3 Particularities Regarding Analysis

> *Economists have inherited from the physical sciences the myth that scientific inference is objective, and free of personal prejudice. This is utter nonsense.*
>
> *Leamer (1983)*

In this section we draw attention to the multitudes of potential problems arising in the field of empirical deterrence research and how they are addressed. However, we will not discuss many of the usual problems arising in empirical studies like, for example, co-integration, unit roots, etc. and how they are treated. Most studies can be categorized into the categories data analysis, surveys and experiments and apply to one nation (or a subset of it).

Studies across multiple nations have to cope with several additional problems. Fajnzylber et al. (1998) point out that many offenses are defined very differently across nations. They resort to homicide which should be the most homogenous crime. In a study of 40 countries from 1970 to 1994 they use five year averages (panel data) or 25 year averages (cross section) because the data is very heterogenous and contains missing data.

Besides the analytical problems addressed in subsection 2.3.1, it is quite disappointing that it still seems necessary to state that "simply finding a correlation between two variables is no guide to whether there is any relationship between the two. In the social sciences it is essential to avoid such naïve comparisons as they are generally misleading and unreliable." (Denny et al., 2004). Such studies should not be underestimated, as Hashem Dezhbakhsh points out (quoted in Donohue and Wolfers (2005)): "the academic survival of a flawed study may not be of much consequence. But, unfortunately, the ill-effects of a bad policy, influenced by flawed research, may hurt generations". Independent of the kind of analysis, it is obvious that each study may contain multitudes of further potential problems of which most cannot be fully solved. Some of these may arise from "bad" data, nasty error terms, model misspecification or unknown independencies. Yunker (1976) emphasizes that "if potential, hypothetical problems like these are taken *too* seriously, they would effectively abrogate the possibility of any kind of statistical investigation, because one or another of them could apply to practically any imaginable project." We agree with Yunker and his conclusion that "awareness of these problems is certainly desirable as a means of avoiding *blind* reliance on statistical evidence that is possibly misleading, but at the same time it would be highly dubious to disregard entirely apparently very strong statistical evidence just because one or more of them might conceivably apply".

Aside from the common ways of analyzing crime data, there are some unusual approaches. Among these, Viscusi (1986) studies crime by analyzing the risk-rewards trade-off, utilizing the existence of compensating differentials. Finding that the premium makes up to two-thirds of the whole crime income, he concludes that "the results bolster the findings of other studies supporting the empirical importance of criminal deterrence". Another interesting model - derived from the principal agent theory - is introduced by Cohen (1987) who remarks that, in some cases, offenses happen unintentional. While he regards environmental offenses, other fields are also possible (e.g., tax evasion). Under optimal conditions no monitoring costs and optimal fines are the first best solutions.

2.3.1 Data Analysis

> *The government are very keen on amassing statistics. They collect them, add them, raise them to the nth power, take the cube root and prepare wonderful diagrams. But you must never forget that every one of these figures comes in the first instance from the village watchman, who just puts down what he damn pleases.*
>
> *Anonymous English Judge, quoted by Sir Josiah Stamp in* Some Economic Matters in Modern Life*, 1929*

Many data sets, especially those providing official rates, offer several measures concerning the probability of detection and punishment: arrest or clearance rates, prosecution rates, conviction rates, police budgets or manpower, and others. Variables which represent the severity of punishment are found more sparsely. In most cases, the average sentence length (for actually convicted individuals or served by those released) is used. When these variables are missing, unusable or undesired, several alternatives have been implemented (e.g., execution rates, license revocation, and many other).

Most studies can be categorized by three classes: cross section, time series and panel data. Although all three classes can be used to study deterrence, Entorf and Spengler (2002) emphasize that “when assessing the impact, it is mainly the cross sectional dimension and not the time dimension from which deeper insights might be expected” and conclude that “it is the appropriate use of panel data that copes with problems arising in empirical crime research”.

In principle, each committed offense is the starting point of a cascade.

1. An offense may be reported to, or detected by the police. The probability that an offense is recorded by the police depends mainly on its type.
2. If reported, the offense may be cleared by arresting or identifying the offender. Again, the probability varies largely by the type of offense (e.g., it is usually low for theft and high for homicide). For several offenses (e.g., shop lifting) reporting and clearing is closely linked.
3. If identified, the offender may be prosecuted (if there is sufficient evidence and the prosecution does not dismiss the action).
4. Conditional upon prosecution the offender might be convicted.
5. Conditional upon conviction, the offender might have to pay a fine, get a probation or be sent to prison. The kind of sentence depends on the type of offense and on the criminal history of the offender.
6. If sentenced to a term in prison, the offender has to spend the whole time or just a part of it in prison.

Because each step is conditioned on the step before, theory tells us that the marginal deterrent effect is reduced in each step. Therefore, an increase of one percent in the arrest rate deters more than an increase of one percent in the imprisonment rate (Wolpin, 1978b). Naturally, any analysis depends on the quality of the available data. In regard to studies about crime in the USA, the NCR is the most commonly used source. Although the Federal Bureau of Investigation (FBI) takes great care to publish reliable and consistent data, their sources remain the individual police departments. And, as Witt and Witte (2000) point out: “at the level of the individual police department, both administrative and political changes can lead to abnormalities in reported data or to failures to report any data.” Especially old data (i.e., older than 1960) may be inappropriate to study deterrence (Yunker, 1976). It is common (at least in the USA and Germany) to associate several offenses committed by a person at the same time (e.g., burglary and homicide) only with the most severe offense. This leads to an underestimation of lesser offenses (Becsi, 1999; Spengler, 2004). Furthermore, it may happen that one crime is committed by several people (e.g., a motor vehicle theft by four people) resulting in several arrests, convictions and penalties for one crime (Braithwaite and Makkai, 1991).

Some authors argue that the usual rates to measure the probability of detection and punishment are not the proper or optimal measures to use. Cloninger (1994) argues that the police presence, measured by the quotient of police and the number of violent crimes, is superior to such rates. Some countries require other measures. For example, Italy is known for its slow prosecution and organized crime in the south. For this reason, Buonanno (2003) uses a measure for the celerity of punishment which is fitted to the Italian scenario and the non-clearance rate. Deterrent effects can be found for the north of Italy while the south is much less affected because "organized crime considerably reduces the efficiency of criminal justice and effectiveness of police force". Organized crime is also used by Marselli and Vannini (1997) to explain insignificant effects of the sentence length (successors and quarrels between them). Tittle and Rowe (1974) suggest that arrest rates (and other rates) may only be effective above a certain threshold value (0.3).

The severity of punishment may also pose a problem. As already mentioned, probation (see section 2.1) may operate in the opposite direction than a prison sentence. However, even the empirical effect of imprisonment may depend on the definition of that variable (Avio and Clark, 1978). Many studies use the (mean) imprisonment length of the released inmates; but these lengths do not reflect the actual but the past severity. Furthermore, the mean sentence length may be biased if the number of convicted criminals changes in each period, even if the distribution of punishment remains the same. Avio and Clark (1978) use the actual mean sentence length and adjust it for parole and remission, and conclude that it is superior to the usual measures but the support for deterrence remains weak. In a cross-sectional study Pogue (1975) uses the difference between the sentences of the local U.S. district court and the overall average of all district courts. Another approach to the severity of punishment is taken by Kessler and Levitt (1998) who employ a difference-in-difference model. They use an increase in the level of punishment for severe offenses in California (the treatment group is defined by the severe offenses) and compare it with the non-severe offenses in California. By subtracting the difference for the United States they conclude that, in the first years, the effect is solely based on deterrence.

Spelman (2000) scrutinizes the prison population and argues that the large decline in crime in the USA in the nineties is only partially caused by the massive expansion in prison capacities and prison population at that time. He argues that this expansion was inefficient because it was responsible for only 25% of the decline in crime. van Tulder and van der Torre (1999) come to contrary conclusions using data from the Netherlands. They find that spending in prison infrastructure is more efficient than spending more on police since investment in the latter does only marginally influence the clear-up rate. Donohue and Levitt (2000) propose a different explanation of the large decline of crime in the USA in the nineties: abortion was legalized in 1973, which is supposed to have led to a large decrease ($15-25$ years later) of children who grew up in problematic surroundings. Levitt (2004) cynically comments on the approaches to explain this drop: "although experts failed to anticipate the decline, there has been no shortage of hypotheses to explain the drop in crime after the fact". Marvell and Moody (1994) study λ (the crime rate of active offenders) and the influence of prison population with several lags. Because prison population deters only up to a lag of one year, they argue that only an incapacitating effect is at work but not deterrence.

In the case that official institutions work on their limit, some rates will be overestimated (Chambouleyron and Willington, 1998). When the legal system is overburdened, a higher arrest rate will, ceteris paribus, imply a lower conviction rate[19]. They try to avoid this bias by using rates per capita as covariates (therefore the denominators are not influenced by other deterrence measures) and perform one regression per deterrence variable; see section 2.3.1 for more information about this measurement bias. Meera and Jayakumar (1995) use this reasoning to argue that the positive effect of prison overcrowding in Malaysia results from the diminishing probability of imprisonment. Bodman and Maultby (1997) also shortly touch this problem in the case of Australia but find only weak evidence that more crime leads to a lower efficiency of the legal system (while resources remain constant).

Using sophisticated methods of analysis does not guarantee good results. The statistical model and its analysis should fit the data to make sound conclusions possible. Studying burglaries in England and Wales, Deadman (2003) observes that the forecasting model which does not rely on distinguishing between short- and long-run effects does achieve the best results. He concludes that there seems to be no long-run equilibrium in regard to the studied time series of burglaries. On the other hand, Sridharan et al. (2003) conclude in their study that "results using regression approaches are biased and the measured effects are not reliable because of the serially correlated errors". Because of seasonal-, trend- and random effects, they prefer Autoregressive Integrated Moving Average (ARIMA) and structured time series models instead of "simple" regression models. Entorf (1996) finds that simple OLS leads to more pronounced deterrent effects than more sophisticated estimators (general IV, error corrected models).

A very important issue of the official crime rates is their reliability. There are several reasons why official statistics may inherit systematic anomalies. Crime rates rely heavily on the readiness of the population to report crimes. Comparisons of victimization reports and official statistics reveal different deviations for most offenses. The best fit is usually to be found for homicide (if detected) and motor vehicle theft (reporting is required to get a compensation from the insurance). Very large gaps can be found for lesser offenses and those which only have a low clearance probability. For example, in the USA in 2004, only 41.4% of all crimes were reported to the police; 49.9% of all violent and 39% of all property crimes. The highest reporting rate (excluding homicide) is found for completed motor vehicle theft (94.8%) while the lowest for attempted purse snatching (17.5%) and completed theft below 50$ (18.8%) (Pastore and Maguire, 2004). Soares (2004) compares international victimization reports and crime statistics and finds that the

[19] When crime remains constant, the decreasing conviction rate will bias its deterrent effect upwards.

reporting behavior is mainly influenced by the grade of development of a nation (measured by the GDP). People in more developed nations more readily report a crime but remain unaffected by their criminal behavior.

All in all, there is a multitude of potential problems, such as those sketched above and described in the following subsections. Some authors try to consider them, some do not. Some simply use OLS and hope that all biases will (hopefully) cancel each other out or are sufficiently small.

Aggregation Bias

Aggregation can occur over the crime categories and (or) over the studied units.

Ideally, every deterrence and crime variable refers to one specific type of crime. In practice, however, this cannot always be done. Many data sets do only contain information about aggregated types of crime on various levels. Some aggregations are generally assumed to be harmless (at least no none complains about it) like not distinguishing various robberies (e.g., of defenseless people, female victims, with firearms, etc.). Other are more problematic, like combining murder with manslaughter, armed with unarmed robberies or even merging whole categories. The latter is studied by Cherry and List (2002) who analyze 70 U.S. counties in the eighties using Index I crimes and their aggregation to property and violent crimes. They conclude that this aggregation leads to unacceptable distortions. Avio and Clark (1976) also emphasize that such aggregations "may lead to unjustified generalizations about individual crime types, and in fact may invalidate such studies as a legitimate attempt to subject the economic model of crime to empirical verification". However, disaggregation (of property crimes) is not an ultimate goal because the offender cannot be sure in all cases how the police (or the judge) will classify the planned crime (Heineke, 1978b) - or that the crime evolves as planned. For example, a burglary may be classified as a larceny or a planned petty theft may escalate to a robbery.

The *classic* aggregation bias goes back to Theil (1954) and is the deviation of the macro parameters from the average of the corresponding micro parameters. However, the question whether or not micro level data (i.e., data about individuals) is superior to aggregated data (e.g., county-, state- or country data) cannot be unambiguously answered. Although the deterrence theory is based on individual responses to incentives, some authors (like Decker and Kohfeld (1990a) or Nagin (1978)) argue that deterrence is meant to influence society itself or, as Nagin (1978) puts it: "general deterrence is inherently an aggregate phenomenon since it is reflected in the behavior of the entire population". However, this may cause some statistical problems because "an equation that holds true for an individual can also be applied to a county, state or nation, only if the functional form is invariant to aggregation" (Dezhbakhsh et al., 2001). This is not the case for log or double-log equations (the sum of log-equations is not another log-equation). Ehrlich (1973) points out that the individual response of offenders to deterrent measures is inhomogeneous because it may vary by their grade of specialization. However, even if aggregation is applicable, there are large differences in the possible levels of aggregation. Lott and Mustard (1997) emphasize that "the very different results between state and county-level data should make us very cautious in aggregating crime data and would imply that the data should remain as disaggregated as possible". At least, failing to incorporate systematic differences (e.g., rural and urban data) is certainly a potential problem in any such an analysis.

Measurement Bias

A measurement bias (also called reporting bias) occurs when the implemented probability does not reflect the "true" probability. It is already addressed in very early studies like Ehrlich (1972). As Pudney et al. (2000a) point out, the measurement error has a random and a systematic component. The random component refers to the prosecution of offenses. Whether the offender is arrested, convicted or sentenced can be seen as a Bernoulli trial. Therefore, even if all crimes are perfectly recorded, the according rates will be binomially distributed around the true values. A systematic error may be introduced by the reporting behavior of the victims or witnesses (under-reporting) and can bias any calculated coefficients. Shifting focus of the police, arrest characteristics, plea bargaining and congestion of the system may pose additional problems (Nagin, 1978).

Depending on the type of offense, the reporting behavior of the population varies by a large margin. On the one hand, many crimes may happen unnoticed: stolen goods are not missed, unsuccessful burglaries not detected, offenses not recognized to be illegal, and similar cases. Even if an offense is witnessed or detected as such, victims might not report it because the costs to report the crime ("waste of time") exceed the expected psychological or materialistic gain; e.g., due to the low probability to solve the crime (Avio and Clark, 1976). The smallest difference between the true and reported crime rates should be found for murder (because of the severity) and auto vehicle theft (because a stolen car has to be reported to get any money from the insurance). However, since an offender cannot know for sure - at least in most cases - whether his crime will be detected and reported to the police or not, the (deterrent) effects on reported crimes should be similar to those which remain unreported (Levitt, 1997b).

According to Denny et al. (2004), as long as the reporting behavior does not vary systematically, or is correlated with other regressors, this does not pose any problems and only adds to the random error term. Or, as Levitt (1998b) states it in more detail: "as long as crime is the left-hand side variable in the analysis, *random* measurement error will increase the standard error of the estimates, but will not bias the parameter estimates. Only measurement error in reported crime rates that is systematically related to the policy being evaluated will bias the estimates". However, especially for cross-sectional data, there may be such systematic differences.

Some authors try to circumvent this bias by adjusting the official rates. Myers (1982) uses victimization reports and, basically, divides the official crime rate by the probability that a crime will be reported. His estimates of the deterrent measures remain negative. Alternatively, instead of the usual clearance rate, the self-reported number of crimes can be used as the denominator. This is done by Craig and Heikkila (1989) who find deterrent effects with this rate but not with the usual clearance rate. Goldberg and Nold (1980) even divide the reporting probability into urban and individual partitions and use the reporting probability in addition to the clearance rate. Only the former is significant. Adjusting macro-data with crime rates from victimization reports is also done by, besides other authors, Cohen and Land (1984) and Lee and McCrary (2005).

A ratio bias, often not distinguished from measurement bias, may be introduced when the exogenous and endogenous variables are mixed in one equation. For example, the typical clearance rate includes the number of offenses in its denominator. As pointed out by Nagin (1978) and several other authors (refer to Levitt (1998c); Avio (1988) or Denny et al. (2004) for more detailed information) this can bias the estimates downwards[20]: "if the intensity with which crime reports, clearances, and arrests are manipulated varies either cross-sectionally or over time, then an inverse association will be generated between crime rates and both clearance rates and arrests per crime, even in the absence of any deterrent effects." Concerning the arrest rates, Eide et al. (1994) note that the elasticity becomes more negative when the portion of unreported offenses decreases. In this context, Avio (1988) compares the estimates of the effect of capital punishment. Instead of using one of the usual execution rates (executions divided by convictions or homicides), he uses the sum of executions and commutations, thus avoiding the ratio bias. He concludes that estimates which are based upon the usual rates are considerably more in favor of deterrence. Instead of the usual conviction rate, Funk and Kugler (2003b) use the absolute number of convictions after controlling for lagged crime (thus they analyze the change in convictions at a given crime level). It is quite common to lag the explanatory variable, as Levitt (1998c,a); Entorf and Winker (2002) or Bedard and Helland (2000) do. Furthermore, it is likely that reporting errors are correlated with education, unemployment, income inequality, etc. (Fajnzylber et al., 1998) which are included as covariates in most empirical analyses of deterrence.

Chambouleyron and Willington (1998) argue that using the clearance-, conviction- and imprisonment rate in one equation gives the correct coefficient of the imprisonment rate but not the correct coefficient of the clearance- and conviction rate (since they share the same terms). Therefore, they replace the imprisonment rate by the imprisonments per capita and re-estimate the equation and take the conviction-coefficient to be the true one. In a last step, they also replace the conviction rate by the convictions per capita and estimate a third time to calculate the true coefficient of the clearance effect.

Pudney et al. (2000a) as well as Levitt (1998c) conclude that the measurement bias is not relevant in practice. Contrarily, Cherry (1999) reports that measurement bias can lead to a gross overestimation of deterrence. He compares an U.S. city panel (using fixed and random effects) to pooled regressions and detects no bias for homicide, rape and motor vehicle theft but a large bias for all other offenses. The largest bias of 70% is found for burglary. Many authors, like Fujii and Mak (1979), simply explain wrongly signed results by dominating reporting effects.

[20] In fact "more sophisticated analysis suggests that the direction of the bias depends upon the actual supply elasticity response" (Avio, 1988), but it is usually assumed to be negative in practice. Ehrlich (1973) also argues that effects in both directions may occur.

A completely different aspect is mentioned by Avio and Clark (1976), who note that clearance and conviction rates may vary locally. In rural areas suspects might be arrested only when a conviction seems to be certain (resulting in lower clearances but higher conviction rates), while the contrary may be the case in urban areas (high arrest- but low conviction rates).

Simultaneity Bias

In principle, an increase of the police resources (e.g., manpower, budget or equipment) should lead to a decrease in crime due to more deterrence. While this may be true on the individual level, a feedback effect may mitigate this in an empirical analysis with aggregated data[21] for various reasons.

- In the long run, crime and law enforcement (e.g., police, laws, etc.) affect each other. More law enforcement reduces crime, and less crime may lead to reduced law enforcement - as well as the other way round: more crime leads to an increased demand for protection. This means that, even if a deterrent effect exists, law enforcement may be positively correlated with crime. The same applies to the level of punishment, which may be increased to meet increased levels of threat of crime, resulting from an increased demand of safety and protection (Koskela and Virén, 1994; Ehrlich and Brower, 1987); however, only in rare cases punishments are lessened when crime decreases. Even when the official level of punishment remains the same, judges may impose harsher sentences as a reaction to increased crime rates (Avio and Clark, 1976).
- Law enforcement expenditures are mainly used to “clean up” after crime and have only little in common with deterring future crime (McPheters and Stronge, 1974).
- In order to avoid budget cuts (or reduced increases), the police may want to exaggerate the actual official crime rates to keep the actual demand for police at least on its current level. This can be done by accepting more charges or by splitting some crimes into several categories or by intensifying activities in crime prone areas (Rasmussen et al., 1993).
- Crime deters punishment due congestion of the legal system (Ehrlich and Brower, 1987). As crime increases, but the resources of the police and courts do not, the efficiency of the police, courts and the prison system (e.g., arrest and conviction rates, actual imprisonment lengths) decreases, when they are already working at their limits.
- More police effectiveness (e.g., induced by a larger budget, more officers, etc.) may decrease the number of false arrests. Therefore, the arrest rate may decrease while the conviction rate increases, although real crime levels remain constant (Sandelin and Skogh, 1986).

[21] The similarity to demand and supply functions is noted by Kenkel (1993): interaction effects have to be considered when dealing with macro data but not when analyzing micro data.

Similar to the measurement bias an increase in the police force or budget may also increase the capability to handle and manage reports by the public and to detect more crimes by themselves (Carr-Hill and Stern, 1973). Thus, increased crime rates, accompanied with more police resources, are often associated with a diminished number of unrecorded cases. Indeed, this explanation is used in many cases when the police variables carry the wrong sign (e.g., by Carr-Hill and Stern (1973); Greenwood and Wadycki (1973); Thaler (1977) or Meera and Jayakumar (1995)). Using monthly data from New York City Corman et al. (1987) conclude that "criminal behavior is more sensitive to changes in sanctions than law enforcement agencies are to changes in crime".

The budget argument (maintaining high crime rates to avoid budget cuts) is also often used as an explanation for positive associations between police variables and official crime rates (as Benson et al. (1998) do). However, according to Benson et al. (1992), most studies concentrate on Index I crimes, while police invest most resources in combating Index II crimes (especially drug offenses). This might already explain many inconclusive coefficients of police expenditures or manpower. Besides these police variables, all other variables - which may be altered when society is faced with increased crime rates - may be potentially affected too. Among these are judges or juries in their readiness to convict (Hoenack and Weiler, 1980) or harsher penalties.

There are several ways to mitigate a potential simultaneity bias. Statistical methods may remove such feedback effects (e.g., instrumented estimators) by using some variables which affect the police resources but not the corresponding crime rate. These are then used to estimate the "true" police resources which are then used in the final estimation. The main problem lies in the identification of such instruments of sufficient quality. Weak instruments may render any calculated estimators unusable. This identification issue seems to be the largest problem - refer to Eide et al. (1994) for a more extensive discussion and examples. Wolpin (1980) emphasizes that these restrictions have to be driven by theory. In combination with aggregated data, Trumbull (1989) criticizes 2SLS estimates in general, because the instruments are often not based on theory, in-between variance may be introduced artificially and inhomogeneity of the analyzed units may introduce a bias. Since OLS is more efficient than 2SLS, the latter is only appropriate when simultaneity is shown to be a problem.

One of the most popular studies in this context is certainly Levitt's usage of electoral cycles (Levitt, 1997b) as instruments. He argues that the number of police officers is periodically adjusted just before elections occur. Since these elections are periodical and determined by general logistical reasons, these elections do not influence crime. Levitt then uses these elections as an instrument to estimate police variables. With these instruments he finds significant deterrent effects of the police. This instrumentation has been cited in dozens of other studies but only rarely implemented - Spengler (2004) is an exception but he does not consider it to be useful. This did not come by surprisingly, because McCrary (2002) shows that Levitt's results are all based on a typo in the implemented algorithm (he used the standard deviation instead of its reciprocal as weights). With the corrected version the results are largely insignificant. In a reply, Levitt (2002b) apologizes, replaces and expands his set of instruments with fire men and achieves similar results (same point estimates; however, these are not significant). Nevertheless, his results seem not to be as convincing as he wants them to be. Up to day, his first paper has been cited many times while McCrary's article and Leviit's reply have not (Nilsson (2004) and Klick and Tabarrok (2005) are exceptions). This is at least, from a scientific point of view, disappointing. In another study, Levitt (1996) uses prison overcrowding as an instrument.

Hakim et al. (1978) employ a very simple line of reasoning to circumvent this bias: police expenditures or manpower is only interpretable as a deterrent when it leads to more arrests. Therefore, any deterrent effect is contained in the arrest rate and the police variable explains something else but not deterrence. On the other hand, Goodman (1997) proposes to incorporate police manpower per capita (which has a positive influence) and then to use police density as a deterrence measure (which has the expected negative sign).

In the case of the death penalty, Zimmerman (2004) derives some instruments from public choice theory. Among others, he uses the number of state murders committed by strangers and the proportion of murders which happened under non-felony circumstances, and the proportion of non-white offenders to estimate the arrest and conviction rates. Furthermore, he resorts to indicators about past botched executions and prisoners released from death row.

Feedback effects may also be dealt with by using lagged variables. However, only a few authors consider or even mention the problem that the assumption of strict exogeneity (that the explanatory variables are uncorrelated with each error term at all leads and lags) may not hold. Among these authors Fajnzylber et al. (1998, 2002a,b); Witt and Witte (2000) and Andrienko (2002) incorporate and consider the concept of weak exogeneity (that each error term may be correlated with future leads but not with the current and lagged values of the explanatory variable). Machin and Meghir (2004) argue that the bias in their analysis (violation of strict exogeneity can make some estimators inconsistent or biased) should be negligible, while Neumayer (2003) and Reilly and Witt (1996) simply assume that weak exogeneity holds for some variables.

It may be argued that potential criminals need some time to perceive any changes in the probability of punishment. On the other hand, police resources definitely require some time to adjust to the current crime rates, due to reallocation of budget resources and manpower, recruiting and training of new personal, etc. (Goodman, 1997). Assuming the latter effect is of more importance, many studies employ lagged police variables, because police variables should be much less affected by future crime rates while crime rates should react to past changes in police resources. In this context, Greenberg and Kessler (1982b) argue that the results in a 2SLS estimate with only two points in time may depend heavily on the assumed lag-structure. Using data of 130 U.S. cities, they are able to find positive significant results for police expenditures as well as negative significant results. Alternatively, high frequent data as in Corman and Mocan (2000) or Corman and Mocan (2002) may be used to circumvent simultaneity. They use monthly police data from New York because police manpower cannot be adjusted to crime on a monthly basis. Moreover, several authors, like Goodman (1997), argue that even yearly data may be sufficient because the police and government need at least one year to react to crime rates.

Comparing cross sections and time series may also help to reduce simultaneity bias. Wolpin (1980) observes that "feedback relationships should differ in relative importance over different observations sets. If deterrence relationships did not so vary, bounds on deterrence could be established within a single equations framework." He argues that long run differences in the level of crime should dominate short run fluctuations in cross sections but not in time series. Therefore, estimates of deterrence tend to be underestimated in cross sections and, because congestion should be a greater problem, to be overestimated in time series.

Incapacitation Bias

When a criminal is imprisoned he cannot commit further crimes while locked away (neglecting crimes within the prison). Assuming that he would commit further crimes, if he had not been imprisoned, the crime rate will decrease. This negative effect is not related to general deterrence and should be taken into account when analyzing crime data. If this is not the case, this incapacitation bias will lead to an overestimation of the deterrent effect. This bias can be avoided, for example, by analyzing shocks in the severity of crimes using VAR-models (Funk and Kugler, 2003b). If sentences are usually conditional prison sentences, no incapacitation effect is present. If the mean served prison sentence is sufficiently large, any short term effect cannot be affected by incapacitation[22]. Wolpin (1978b) compares the deterrent effect of two types of sentences imposed on guilty offenders (prison or non-prison sentence). Another possibility is to introduce the concept of imperfect foresight to the model. Only anticipated changes in the deterrent rates can effectively deter while all other effects have to be attributed to incapacitation. Wolpin (1978b) concludes that, for crimes against the person, the incapacitation effect is almost equal to the deterrent effect (in both models of perfect and imperfect foresight).

Levitt (1998a) studies juvenile crime and uses the transition from juvenile to adult courts to distinguish deterrence and incapacitation. Deterrence implies a sharp reduction in the transition while incapacitation implies a smoother transition due to lags in the arrest and imprisonment process, as well as mild sentences in the beginning because juvenile records are sealed after reaching adulthood. The sharp drop found in the data indicates that the incapacitation effect is very small.

In another paper, Levitt (1998c) uses cross-crime effects to isolate deterrence from incapacitation. When the arrest rates for one crime increase, deterrence predicts an increase in all other crimes because criminals will switch to other, relatively less deterred crimes (for crime switching see also subsection 2.2.7). In contrast, incapacitation predicts a decrease in all crimes, since the number of available offenders is reduced. He finds that deterrence is more relevant than incapacitation.

Levitt (1996) supplies evidence that the costs of a prisoner is of the same magnitude as the social harm an offender causes. However, Holtmann and Yap (1978) point out that the relative

[22] The improved incapacitation effect of a sudden increase in the average sentence length becomes relevant not before the old average sentence length is surpassed.

costs of imprisonment for theft is too high when compared with those of robbery and burglary because the loss in the case of theft is usually quite low.

Misspecification Bias

Model uncertainty is composed of at least two parts: sampling uncertainty and specification uncertainty. While the former ever decreases as sample size increases, the latter remains constant (Leamer, 1983). Including unimportant variables or omitting variables which are not correlated with other explanatory variables do just inflate the standard deviations, but omitting important variables can lead to a systematic bias. While the former does just decrease the predictive power of the model, Entorf and Spengler (2002) note on the latter "that the higher the influence of the omitted variable on the explained variable and the higher the correlation between the included and the excluded variable are, the higher will be the omitted variable bias of the estimated coefficient of the included variable". While the number of observations is usually limited the number of applicable covariates is not.

Mustard (2003) studies misspecification effects for the conviction rate and the sentence served, while keeping the arrest rate in the equation. He uses panel data from four U.S. states on county level. Because the arrest rate is negatively correlated with the other two variables (the conviction rates and sentenced served of the marginal offender decreases when the clearance rate increases), effects of the arrest rate are underestimated up to 50% when omitting such important variables (e.g., the elasticity of the arrest rate for auto theft expands from -0.0027 to -0.0052 when conviction rate and sentenced served are added to the model). Gyimah-Brempong (1986) reports, analyzing all cities in Florida, that not distinguishing important social and economic variables by race leads to a large bias for the minority. For example, the unemployment rate is dominated by whites, and non-whites are found to be more prone to crime. However, when using the white and non-white unemployment rates, the non-white dummy becomes insignificant and changes its sign.

Ruhm (1996) studies the effect of omitted variables for DUI offenses and finds that omitting the effect of organizations like MADD, or factors like beer taxes, lead to an overestimation of the effects of anti-DUI laws. Similarly, Entorf and Winker (2002) (using German data) and Benson et al. (1992) (using U.S. data) argue that ignoring drug consumption (at least in times with high drug consumption) will lead to misspecified models and biased results. However, while the marginal deterrent effects may be altered, their signs seem to remain correct (Eide et al., 1994).

In general, not considering other important variables (when they are correlated with included variables) can bias the estimates in any direction. Pogue (1986) argues that the results, e.g., those of Ehrlich (1973), are overestimated because "it is possible to obtain statistically and quantitatively significant crime prevention effects by selecting a particular equation specification, for example, one that includes relatively few exogenous variables, or a particular cross-section year". As described in section 3.4, any desired result may be achieved when the specification is chosen accordingly. Especially among natural- and quasi experiments, as well as in the analysis of crime data, it is quite common to use dummy variables as the relevant deterrence variables. However, Dezhbakhsh and Rubin (1998) point out that any analysis based only on a regime shift effect, measured by a dummy, may be biased when other regressors also correlate with the dummy. Leamer gives a nice and simple example of deriving different conclusions from the same data set:

> *The applied econometrician is like a farmer who notices that the yield is somewhat higher under trees where birds roost, and he uses this as evidence that bird droppings increase yields. However, when he presents this finding at the annual meeting of the American Ecological Association, another farmer in the audience objects that he used the same data but came up with the conclusion that moderate amounts of shade increase yields.*
>
> *Leamer (1983)*

To minimize model misspecification, many authors analyze numerous specifications to test whether their favorite specification remains unaffected; or simply to find the "best" specification according to a chosen statistic (e.g., van Tulder and van der Torre (1999) use the R^2). Other authors employ data mining methods, as described in more detail in section 3.6, like Extreme Bounds Analysis (EBA) or Bayesian Model Averaging (BMA) to extract robust estimators.

Replacement Effect

In some cases local efforts to combat crime do only redistribute crime to other places. This effect is commonly called replacement or spillover effect. For example, using camera or lightning on specific crime prone places or intensifying patrols in certain areas may just displace crimes to lesser monitored sites.

Mehay (1977) studies the Los Angeles metropolitan area using the differences in patrol intensity to detect such spillover effects. Although he finds such effects, these are only small and of minor importance. Fabrikant (1979) hypothesizes that the criminals spatial choice depends on economic gain and competitive pressure and he finds that juveniles prefer to commit offenses in their own districts. This is also supported by Chaiken et al. (1974) and Farley and Hansel (1981) who report that delinquents tend to commit their offenses near their homes (with the exception of rape), and that the relative deprivation and central city decline is more important than the relation between the city and metropolitan population. Data of the metropolitan area of Montreal in Canada (Furlong and Mehay, 1981) also seem to indicate such spillover effects. These may bias results when estimating deterrent effects of the police. Additionally, replacement effects may induce more spending on law enforcement than would be necessary without such effects (Rasmussen et al., 1993). However, Press (1971) does not find such displacement effects. He studies an increase of policemen on the street in the 20th district of New York and finds that crimes (visible from the street) were significantly reduced but no effects could be observed in the neighboring districts.

Similar to replacement effects is another phenomenon. Even if crime is not pushed out, negative effects can occur because people tend to move away from areas with higher crime rates if they can afford it. This may lead to increasing crime rates because the remaining population is more prone to crime; such a vicious circle may accelerate the neighborhood decline (Katzman, 1980). Burnell (1988) finds that crime reduces the house values; not even for the affected area but also for neighboring municipalities. Clark and Cosgrove (1990), studying willingness-to-pay for public safety, also find that crime affects land values. People in the more crime prone central city are willing to pay less for public safety than those living in safer suburban areas.

Wealth effects are also associated with the distribution of crime. Hakim (1980) studies the metropolitan areas of Camden and Philadelphia, and reports that wealthier cities spend more money on police but also attract more criminals from nearby areas with a good traffic connection. The overall effect is that crime increases for those cities, although police expenditures are increased as well. A similar effect is observed for areas with many tourists like Hawaii. Fujii and Mak (1979) analyze the agricultural displacement in Hawaii (changes of employment in the agricultural and hotel sectors) and report that tourism significantly fosters property crime and rapes (tourists bear twice the risk than residents to become a victim). Local displacement is also found for security appliances (e.g., alarm devices) when the number of potential victims is large enough (Ayres and Levitt, 1998). Clotfelter (1978) adds that private security measures may also cause a replacement effect, diverting crime to lesser protected victims or houses. He indicates "that the greater the relative importance of such a displacement effect, the more private protection will tend to be oversupplied, from a social viewpoint". Furthermore, there may be an isolation effect when people avoid, in fear of victimization, locations they would have visited otherwise.

Beside the spatial component, Chaiken et al. (1974) also study replacement in time. They analyze a large increase in police presence in subways during the night in New York. While there is a significant deterrent effect during the night, two other effects are also observed: a phantom effect during the day (the number of offenses decreased although police presence remained constant), and an increase of offenses in buses. After eight months, the deterrent effects had faded and crime increased significantly. They conclude that this "tends to confirm that potential offenders do in fact try to estimate the risks of criminal activity".

Short- and Long Term Effects

While many studies do not - or simply cannot - distinguish between immediate effects and effects in the long run, some do. There are especially two analytical approaches: VAR- and error-correcting models (ECM). These approaches are used by several authors and belong to the common tools of advanced statisticians.

VAR-models are used to determine the effect of a shock in one variable in a system over time. Witt and Witte (2000) study the effect of the prison population on crime and find that the number of prisoners (serving more than one year) has no short-term but negative long-term effects on the crime rate. Corman et al. (1987) emphasize that standard time series analyses suffer from weak identification restrictions which may be circumvented by using VAR-models. Funk and Kugler

(2003b) argue that the "major advantage of VAR estimations, compared to traditional panel data or cross-sectional analyses, lies in the better understanding of the crime-reducing impact of harsher governmental enforcement".

Error correcting models are usually estimated with standard regression methods like OLS. However, difference operators and lag structures can be used to distinguish a long term equilibrium from short termed deviations. Entorf and Spengler (2000) study crime in western Germany, using static and dynamic (ECM) panel estimators. The short and long termed coefficients are quite similar, "indicating a relatively quick convergence to equilibrium", and support the deterrence theory. Pyle and Deadman (1994) emphasize that "incorporating error-correction mechanisms, highlight the need to build a convincing dynamic model of criminal activity and its relationship to the economy as a whole". They study time series of robbery, burglary and theft in England and Wales and find that there is a significant tendency of crime to "bounce back in the long term, so that the long-run equilibrium relationship is restored".

Another aspect is the time frame between the actual change in deterrence and the perceived change. Studies using yearly data usually use the crime rates and deterrence variables of the same year. However, this official information is usually not available until, at least, the following year[23]. However, the "enforcements measure to be used in the regression equations should reflect the expected sanctions, not the level of sanctions observed after the deterrence effect has operated" (Magat and Viscusi, 1990). It is not obvious how fast potential offenders adjust their perception of deterrence. Eide et al. (1994) argue that it is more likely that potential offenders base their perceptions on several years. In contrast, authors like Corman and Mocan (2000) use high-frequency data (e.g., monthly data) because potential offenders can adjust their perceptions in such short time frames but institutional changes (e.g., the number of employed police officers) require more time. Nevertheless, it is unclear how fast offender really adjust their perceptions and, consequently, some authors utilize lag-structures which optimize their model by some criteria, as, for example, Masih and Masih (1996) or Corman and Mocan (2000) do.

Functional Form

Although there is a huge variety of functional forms implemented to model crime, two stand out: the simple additive and multiplicative form. Trumbull (1989) argues that the functional form should be chosen which optimizes the model quality (using Box-Cox-Transformations) while Hoenack et al. (1978) state that the "econometric structure should be specified in the mathematical form which conforms most closely to behavioral expectations". Ehrlich (1977a) emphasizes the multiplicative relationship between the probability and severity of punishment in the case of capital punishment and Mendes and McDonald (2001) for all offenses. Mendes (2004) shows that when decomposing the logarithmic rates the deterrent effects of these factors are almost equal. She concludes that the usual assumption, that criminals are risk preferrers (i.e., the probability is

[23] For these studies subjective measures are usually unavailable anyway.

more influential than the severity), may be just an artifact of the implemented functional form. Furthermore, disregarding the multiplicative structure bears an intrinsic weakness when dealing with very low detection probabilities. Even Marchese Beccaria (1819) noted that very low detection probabilities had to be compensated by more severe punishment. However, the probability and severity of punishment affect crimes independently in an additive model. Stafford et al. (1986) concentrate on the difference between an additive and interactive model. They compare additive, matching and satisfaction balance models using U.S. homicide data and experiments. In all cases the interactive model explains more variance; and punishment deters more (i.e., is more significant) if the variance of a model decreases.

Unobserved Heterogeneity

When comparing different sets of objects (for example, counties or states), it seems reasonable to assume that there are factors which systematically vary among these objects and influence the outcome under study but cannot be accounted for. When these effects are constant in time, fixed or random effects absorb these influences in a panel data analysis. However, cross-sectional data may be severely affected by unobserved heterogeneity. Not considering these effects can considerably bias the estimates. Cornwell and Trumbull (1994) show that the estimated elasticities can be reduced by 30% when unobserved heterogeneity is considered. Unobserved influences may also vary in time (e.g., common shocks influencing all units differently). Additionally, these influences do not have to be stationary; refer to Coakley et al. (2006) for more information about these problems and estimators which are employed in these cases. However, time-varying and not stationary unobserved heterogeneity are not considered in the empirical deterrence literature.

2.3.2 Experiments and Surveys

> *An experiment is like a radio: if we twiddle the knobs at random, there's no telling what we will find, nor any guarantee that it will be in a language we understand, even though the radio itself may be in perfect working order. On the other hand, if the radio is accurately tuned, we can expect to hear something, and also, which is especially important, we can expect others whose radios are similarly tuned to hear the same thing.*
>
> *Eiser (1986)*

Aggregated crime data is, naturally, not usable to test the deterrence hypothesis on the individual level. Therefore, individual data are usually acquired in two different ways: conducting laboratory experiments or using surveys. The big advantage that these data may provide is the possibility to make use of the subjective perceptions of the probability and severity of punishment. Such data may be more close to the theory of deterrence, while it might be less useful to draw conclusions for public policy (see section 2.3.1).

Classic experiments commonly take place in controlled surroundings, such as games or treatments, while quasi and natural experiments usually exploit changes in the environment, which are assumed to be exogenous, and use aggregate data (e.g., law changes). The experiment which may be closest to a classic medical experiment was done by Sullivan et al. (2001). The memory of four groups were tested (RAVLT-test) who simulated to be affected by a disease lowering the memory performance. Two groups were given an abstract warning that any simulation could be detected somehow, but the test results did not show any deterrent effect of that warning. Most laboratory experiments, conducted as games, can be found in the literature about tax evasion and deterrence; refer to subsection 2.2.4 for more references and information.

Experiments in the field are also conducted. Press (1971) studies an increase of 40% of policemen on the streets in the 20th district of New York. He detects significant decreases in "outer crimes" which are visible from the streets; indoor crimes were not affected. Laycock (1991) made an experiment in three villages in South Wales in which some inhabitants marked their property and announced this to the public (e.g., using clearly visible signs on the window or door). Indeed, the burglary rate significantly decreased in the following two years for the participants but not for the other villagers. Chaiken et al. (1974) study a large increase of police presence in the New York subways during the night and find a strong deterrent effect (even during the day) which lasted for eight months. In the years 1976-78 special cars were used in the city of Stockton to detect drunken drivers during the night at weekends (Voas and Hause, 1987). BAC levels and the number of accidents decreased during all nights but not during the day. Retting et al. (1999) study the effect of red light cameras in Oxnard. These cameras reduced the crossing of red lights not only for those stoplights with cameras, but also those nearby (while no effect was detected in nearby cities without such cameras). Another traffic experiment was done by Michaels (1960) who studies the patrol density in Wisconsin on several routes (4 experimental and 3 control routes). The only effect he finds is a reduction in the speed-variance but no other deterrent effect is detected (in regard to the average speed, speeding and accidents). Buikhuisen (1974) reports an experiment in two Dutch cities. The percentage of worn tires was measured in both cities but only in one city it was made public that the police was looking especially for worn tires. Indeed, only in the experimental city the percentage of worn tires was significantly lowered.

Many political decisions also provide good opportunities to test certain hypotheses. For example, Hansen and Machin (2002) use the introduction of minimum wages in the United Kingdom to study the relationship between wages and crime. Indeed, they find that crime is significantly reduced in areas in which very low wages were common. Other quasi-experiments are commonly found in the literature about drunk driving (see subsection 2.2.6), which study the impact of new laws against DUI activities.

Authors using survey data usually use public surveys, like the NCVS or NLSY in the case of the USA, or conduct their own surveys. Some of these are representative, some are not and often this status is simply undefined. These surveys can be sorted into three categories: cross-sectional, longitudinal (multiple cross sections) and panel surveys. Although cross-sectional surveys have been used in many studies, these are criticized to be unable to distinguish the deterrent effect from an experiential effect (see, for example, the studies of Saltzman et al. (1982); Minor and Harry (1982) and Bishop (1984a) who use two-wave surveys, and a discussion on this problem). Surveys are, besides experiments, very prominent in the research of tax evasion and in the sociological and criminological literature.

Lochner (2001) develops a model with a bayesian belief updating system. Offenders adjust their subjective probability upwards if they are caught, and downwards if their deeds remain undetected. He tests his model with data from the NLSY97 and National Youth Survey (NYS). He finds that the real probabilities are grossly overestimated. The adjustment process of the clearance rate and the severity of punishment takes several years. Real and subjective rates are not correlated for youths and then slowly converge towards the official rates. In another study about his belief updating system, Lochner (2003) concludes that the individual perception of the probability of arrest is hard to explain but is influenced by prior arrests (of oneself or friends), the general clearance rate, age and other factors. While the actual probabilities are generally overestimated the perceived order is correct.

Using a survey of 1700 pupils from Arizona, Erickson et al. (1977) conclude that the subjective severeness of offenses is the decisive factor which dominates the subjective probability and severity of punishment. In fact, the probability and severity are correlated so much that they cannot be disentangled. Thus, not negative but positive deterrence - the preventive effect of the normative attitude - is found to be the decisive factor.

2.3.3 Social, Human and Criminal Capital

> *The affluence of the rich excites the indignation of the poor, who are often both driven by want, and prompted by envy, to invade his possessions. It is only under the shelter of the civil magistrate that the owner of that valuable property, which is acquired by the labour of many years, or perhaps of many successive generations, can sleep a single night in security. [...] Where there is no property, or at least none that exceeds the value of two or three days labour, civil government is not necessary.*
>
> *Smith (1789)*

In principal, social, human and criminal capital represent the knowledge and skills accumulated by individuals in these categories. For example, Williams and Sickles (2002) use education and social ties to measure social capital. They find that accumulating social capital reduces (property) crime; i.e., the potential loss of social capital deters. Gyimah-Brempong (1986) emphasizes that criminal capital, in case of the USA, has to be distinguished, at least for the white and non-white population, because general social and economic measures cannot achieve this.

Leung (2002) studies social, human and criminal capital of lower-class youths in Montreal. He employs logit models to identify the influence of education, work, friends and family and general deterrents (police expenditures) on self reported delinquency. While police is insignificant, motivation in school and living with two parents are negatively, while having delinquent friends are positively correlated with delinquency.

Although juvenile law is more lenient, the expected punishment for adults may reduce juvenile crime because, due to dynamic deterrence, these punishments reduce the expected human capital in the future (Levitt, 1998a). One consequence is that harsher punishments may reduce crimes in the short run but increase crime in the long run since criminal capital becomes more important relative to human capital.

Recidivism might be interpreted as failure of special deterrence. However, Ehrlich (1972) remarks that "it might rather be the result of choice dictated by opportunities". Being in prison increases criminal capital[24], while decreasing human capital (e.g., lowering future income). Therefore, it is not surprising when the individual optimal participation in crime remains unchanged or even tips towards crime. Whether or not the loss in human and social capital exceeds the gain from increased criminal capital depends on the individual case. More on that topic can be found in Meyer (2007) who studies social capital and delinquency for Germany using a large survey of prisoners and non-incapacitated people.

A model of human capital and crime is developed by Lochner (2004) who considers the skill to learn, income and the opportunity costs to unlearn (due to imprisonment) and various offenses. Using data from the UCR and NLSY, he concludes that older, intelligent and educated people do commit less simple offenses except for white-collar crimes.

Mocan and Bali (2005) argue that crime behaves acyclically because crime will decrease less when deterrence is increased than crime would increase in the case of decreasing deterrence. This follows from the assumption that offenders build up criminal capital which lessens the efficiency of deterrence. They find evidence for this thesis in the case of property crime but not, as expected, for homicide and rape. In another study, Mocan et al. (2005) develop a two stage model, in which the first step consists of the allocation of time in the legal and illegal sector, taking into account potential income in both fields and deterrence measures. Each individual can accumulate capital in both sectors. In the second step, depending on the earned income, consumption is determined. In the long run, depending on the chosen conditions, the model inhibits one (only legal or illegal activities) or a second (mixture of legal and illegal activities) equilibrium.

More criminal capital may also imply that more experienced criminals will be detected less often or concentrate on crimes which are more inconspicuous (Benson et al., 1992). Crimes, which require more skill and experience may be more affected by conviction rates while other

[24] Crime skills may be trained in prison (Avio and Clark, 1976).

depend more on clearance rates (Avio and Clark, 1976).

2.3.4 The Typical Age Curve

It is the mark of an instructed mind to rest assured with that degree of precision that the nature of the subject admits, and not to seek exactness when only an approximation of the truth is possible.

Aristotle

It is known since Quetelet (1831) and commonly observed that criminal activity rises strongly for young juveniles, reaching its peak around the beginning of adulthood (about the age of 20) and then recedes (Hirschi and Gottfredson, 1983). Hirschi and Gottfredson (1985) even argue that age is a primary variable to explain crime. The observed age curve can be explained by different means. Lochner (2004) uses a model, based on the accumulation of human and criminal capital evolvement (see subsection 2.3.3), and notes that youths commit less simple crimes as soon as they start to work. He presents another explanation in Lochner (2001), in which the curve follows from a model of individual belief updating. Older people do not become criminals because they never adjust their perceived probabilities of arrest downwards, while older criminals adjust their probability upwards until crimes stop to pay off. He adds that unambitious criminals drop out of the criminal market faster than ambitious (and more skilled) offenders. Cohen and Land (1987) argue that age structure, business cycle, criminal opportunity and the imprisonment rate can explain most of the annual homicide and motor vehicle theft rates. Steffensmeier and Harer (1987) report that about 40% of the drop in crime rates in 1980-84 can be explained by age adjustments. However, Greenberg (1985) or Shavit and Rattner (1988) cast doubts on some of these explanations.

There is a large branch of literature dealing with the number of crimes committed by an offender (called λ), which is important when studying criminal careers - and criminal careers depend heavily on the age structure. For further references, see Free Encyclopedia (2008) for a bibliography about criminal careers.

2.4 Particularities Regarding Covariates

There must be no barriers to freedom of inquiry. There is no place for dogma in science. The scientist is free, and must be free to ask any questions, to doubt any assertion, to seek for any evidence, to correct any errors.

Robert Oppenheimer

Hale (1999) points out that “any model of crime trends must include variables which might be considered to capture the deterrence effects of the criminal justice system”. This encompasses a vast pool of possible covariates, of which some are included in most studies if available (like unemployment, race, education and income) while many others are used less often. Pogue (1975) argues that economic factors may also influence criminal behavior indirectly by partially determining the basic attitude towards society and its values. Commonly, the desired measures are not available and are estimated by proxy variables. Any analysis which lacks these influences might be severely biased in any direction (see section 2.3.1). For example, Sridharan et al. (2003) studies the abolishment of parole and increased punishment for felony offenders in Virginia, 1995. The obvious deterrent effects for murder and rape are nullified when unemployment is included in the model. The influential role of the composition of variables is also pointed out by McManus (1985) in his re-evaluation of a study about the death penalty in the USA. His approach is similar to an Extreme Bounds Analysis (see subsection 3.6.4) in such a way that he makes several groups of covariates, according to the hypothesized prior belief of a researcher (from a "right winger" to a "crimes of passion" advocate). He shows that the evaluated results can not be considered robust because the outcomes vary from strong deterrence to absolutely no effect.

In regard to western nations, many economic, social and environmental factors are regarded to be important when studying crime. There are only some rare studies in English which are concerned with African, ex-Soviet or Asian nations. These few studies indicate that their crime-structure differs from western nations by a large margin (Mui and Ali, 1997).

Gross and Hakim (1982) point out that crime characteristics also vary by local transport connections. They analyze the metropolitan area of Philadelphia and find that suburban communities, which are easily accessible, attract offenders from other areas. More detailed results are delivered by Ihlanfeldt (2003), who studies rail stations in Atlanta, Georgia, and finds that the effect of rail stations on crimes is conditional. "Rail access does increase crime within those neighborhoods that are both close to poor people and are not high-income", while rail-transit even reduced crime for the representative white suburban neighborhood. Like Ihlanfeldt, Thaler (1977) finds, using detailed data about city districts of Rochester, that the probability of arrest decreases as the distance between the home of the offender and the site of crime increases.

2.4.1 Unemployment

An optimist thinks this is the best of all worlds. A pessimist fears the same may be true.

Doug Larson

The relationship between crime and unemployment has been subject to numerous studies. Field (1990) reports that studies about crime and unemployment reach back to the 19th century. Although of primitive nature, these studies asserted that economic downturns are positively correlated with property crimes[25]. Only in the last decades, deterrence and unemployment variables are analyzed at the same time. In fact, most of those studies which use deterrence as a covariate,

[25] However, these studies are not included in our data base because they do not provide any measures of deterrence.

focus on unemployment. Deadman and Pyle (2000), studying a long time series of burglary in England and Wales, conclude that unemployment fosters burglaries, while Wong (1995) comes to the same conclusion using crime data of England and Wales from the 19th century. However, these results may depend on the chosen models, because Field (1990) also uses data from England and Wales, but identifies consumption as the primary driving force which renders unemployment irrelevant. A more detailed analysis is done by Hale (1999) who uses an error correction model to distinguish short and long term effects. Studying burglary and theft in England and Wales, he concludes that unemployment and consumption have only short termed effects (unemployment positive and consumption negative). Long term effects are detected for the decreasing size of the manufacturing and producing industry which is explained by the long term increase in the low-skilled unemployment sector.

Corman et al. (1987) use monthly data from New York City and don't find relevant effects of unemployment on property crime. Contrarily, Ralston (1999) uses U.S. time series and detects significant effects of cyclical and frictional unemployment as well as technical unemployment interacting with the arrest rate of whites. Diez-Ticio (2000) emphasizes that the short- and long term effects may be different. He finds no distinctive short term effect but reports that unemployment reduces crime (robbery, burglary and auto vehicle theft) in the long run. Bodman and Maultby (1997) find that long term unemployment leads to more property crimes, while the strength of this effect is smaller for women and short term unemployment. Machin and Meghir (2004) remark that the significance of unemployment on crime vanishes when they introduce spatial fixed effects into their model. Similarly, Field (1990) reports that the effect of unemployment vanishes in the case of property crime when he uses consumption as a covariate.

Weinberg et al. (2002) use individual data from the USA and concentrate on unskilled male workers. While the general unemployment rate and wages have no significant effect on crime, the unemployment and wage development for this subgroup have a significant impact. Especially the long-term wage development is important for the long-term crime progression. Educated men are not affected by these factors. The difference between the desired and actual time of employment of teenagers is used by Good and Pirog-Good (1987) to catch frustration effects. They find that among black teenagers, fewer engage in crime while white teenagers are unaffected and higher police activities only affect whites. This indicates that "blacks apparently view employment and crime as alternative income-generating activities". They argue that reducing unemployment for high-risk black youths additionally reduces crime.

Empirical studies often find that the employment for young workers is significantly reduced when a minimum wage (above the market wage) is introduced (or increased). Since increased unemployment can also increase crime, Chressanthis and Grimes (1990) test this hypothesis[26] with U.S. data. While they find only effects for homicide, rape and motor vehicle theft, Hashimoto (1987) finds evidence supporting this theory for all teenage UCR crime rates.

[26] They also give many references to the associated literature.

Mocan et al. (2005) point out that increasing unemployment may increase crime which implicates the accumulation of criminal capital. When unemployment is reduced, the criminal capital should then dampen the decreasing effect on crime. Similarly, Sherman et al. (1998) argue that special prevention is less effective for unemployed people.

2.4.2 Income, Welfare and Poverty

I know that you believe that you understood what you think I said, but I am not sure you realize that what you heard is not what I meant.

Robert McCloskey, U.S. State Department spokesman

The potential relevance of income influencing crime has been known for about 150 years, as can be seen by von Mayr (1867) who reported a positive relationship between the price of rye (a reverse proxy of real income and consumption) and property crime in Bavaria (as quoted in Field (1990)). In theory, measures of income have two opposite effects. On the one hand, lower levels of income can motivate property crime (as a second income, decreasing opportunity costs) and violent crime (due to frustration). However, it can lower property crime as well because the overall value and availability of loot is reduced and homes are more often occupied (Ensor and Godfrey, 1993). Which effect dominates depends on the wage distribution (Machin and Meghir, 2004). Hence, at least two different variables should be used - such as real income to proxy potential loot and wages to measure opportunity costs of crime (Doyle et al., 1999).

Using Swedish panel data, Nilsson (2004) finds that the social class below 40% of the median income is more prone to burglary and auto vehicle theft, while the average income is negatively related to property crimes (property is better protected). Contrarily, the 90% quantile of the income distribution is positively related to property crime, which is explained by the more valuable and available loot. Machin and Meghir (2004) also conclude that crime rises when the income of the lowest 25% decreases. Danziger and Wheeler (1975) emphasize that, studying time series and cross sections from the USA, combating income divergence is more effective than deterrence in reducing crime. Furthermore, more deterrence can lead to more false convictions and does not reduce the social gap. Myers (1982) also reports that “higher income is a better deterrent to some crimes than increased punishment”. In contrast, Doyle et al. (1999) detect no effect of income inequality but “find that crime is most elastic with respect to wages in sectors that use low-skilled labor”, and that good labor market conditions have a negative effect on crime in general. Using panel data from England and Wales, Witt et al. (1998) use wage differences and unemployment but find only significant effects for shop lifting. In a study of the U.S. county Mahoning, Liu and Bee (1983) implement local data of unemployment and the close-downs of plants. They summarize that local economic variables are important, not nationwide statistics.

Instead of using income, Sesnowitz and Hexter (1982) use the incurred losses reported to insurances in the case of burglaries. They argue that “the present study provides more direct support for the hypothesis that thieves respond to the amount available for stealing". A different approach is taken by Goldberg and Nold (1980) who calculate different burglary probabilities for various values of loot which are then used in the deterrence model. Zhang (1997) studies the effect of welfare-programs on crime and finds that potential offenders who are risk-avers refrain from crimes when welfare payments are stopped in the case of a conviction.

Weinberg et al. (2002) emphasize the importance to distinguish between unskilled and educated men. Using data from the USA, they find that wages and other economic conditions have no effect on (higher) educated men in regard to crime. However, unskilled young men - two thirds of all prison inmates in the USA have no high school graduation - significantly react to changes in wages and unemployment in their delinquent behavior. Viscusi (1986) emphasizes the importance of income even more: the standard approach (regressing deterrence on crime) is only valid if there are no differences in crime income levels. In his analysis, he does not only find significant deterrent effects but also concludes that deterrence is a major determinant of the criminal income.

2.4.3 Education

There is only one good, knowledge, and one evil, ignorance.
Socrates

Education may be associated with more or less crime. On the one hand, it increases human capital as well as the current income and thus increases the opportunity costs of crime. It may also have a *civilization effect* (Usher, 1997) which tends to increase the reluctance to commit an offense. However, education may increase crime as well for several reasons, as Ehrlich (1975b) points out. The marginal product of labor is larger in the illegal than in the legal sector (more criminal capital is accumulated than human capital); higher education may lead to less under-reporting (see also section 2.3.1); education may be a proxy for "the average permanent income in the population, thus reflecting potential gains to be had from crime, especially property crimes"; and some crime rates may be "directly related to inequalities in schooling and on-the-job training".

Besides including education as a covariate in an analysis, like numerous authors do, Lott (1987) explicitly studies whether lower education increases crime. He uses data from U.S. counties to study youth delinquency in regard to the type of attended school (public or private). Indeed, he finds that youths from public schools are more prone to youth related crimes. Using data from England and Wales of the 19th century, Wong (1995) finds that increasing education reduces crime. Bodman and Maultby (1997) point out that effects from education (and similar variables like immigrant- or native status) might affect unemployment directly and crime just indirectly. Fajnzylber et al. (1998) come to the conclusion that "there is a delayed effect of educational effort on crime alleviation" because education does not affect the delinquent behavior of youths, but affects them when they reach adulthood. Additionally, there are indirect effects due to the influence of education on their economic and social status.

2.4.4 Other Variables

Facts do not "speak for themselves". They speak for or against competing theories. Facts divorced from theories or visions are mere isolated curiosities.

Thomas Sowell

Many authors have found that the portion of black people (Dezhbakhsh et al., 2001) or non-whites do significantly increase the homicide rate. However, Pogue (1986) argues that non-white variables might not measure racial effects but are "a fairly good proxy for the frequency of broken homes". Blacks (or non-whites) are often found to be more prone to crime than whites which is often explained by their worse economic status. However, Gyimah-Brempong (1986) reports that, for example, the unemployment rate is dominated by whites and when the unemployment rate is considered separately for both groups, the non-white variable becomes insignificant. Nevertheless, Mocan and Rees (2005) use representative micro data of juveniles in the USA and find significant race effects, even when considering many personal, family and neighborhood characteristics.

Besides for DUI offenses (see subsection 2.2.6) alcohol consumption may increase crime rates. On the one hand, alcohol consumption may increase the probability of detection by impairing the offender (Ensor and Godfrey, 1993). On the other hand, some offenses are often committed under the influence of alcohol (e.g., loitering, assault). The former, however, applies foremost for offenses which require no or little planning like assault or robbery and for these offenses deterrence is assumed to be less effective.

Aside from studies about the death penalty (e.g., to distinguish "harsh" from "lenient" states), political variables are not used that often. While variables for different government constellations are significant in a German panel data (Entorf, 1996), these effects become insignificant when fixed effects are added to the model. Fischer (2004) and Feld and Frey (2004) analyze, aside several other variables, the various levels of democracy of Swiss cantons to identify important influences on crime and tax evasion. However, direct democracy has no effect on crime but increases the tax morale.

It is commonly assumed that crime is more of a problem in cities; i.e., areas with a high population density. However, Howsen and Jarrell (1987) find that there is a U-shaped influence. This means that areas with a very low population density show high crime rates as well. Witt and Witte (2000) use the female labor force participation rate as a proxy for the social development of a society and find that the former is strongly and positively correlated with crime rates, although they "cannot unambiguously say that increases in female labor force participation *cause* crime".

The existence of state lotteries and its influence on crime rates is studied by Mikesell and Pirog-Good (1990) with a U.S. panel. Previously, state lotteries were found to reduce illegal gambling as well as acting "like a regressive tax to the relative detriment of low income individuals". However, they find that having state lotteries increase crime rates significantly by three percent.

2.5 Interim Summary

Justice is conscience, not a personal conscience but the conscience of the whole of humanity. Those who clearly recognize the voice of their own conscience usually recognize also the voice of justice.

Alexander Solzhenitsyn, 1967

All in all, there is clearly an abundance of potential problems, biases, attitudes and theoretical approaches when empirically studying deterrence; in addition to all other usual difficulties analyzing time series, cross sections, surveys, other data or experiments. Brier and Fienberg (1980) summarize some economic studies about deterrence - especially from Ehrlich, Forst and Loftin - and discard all their results. In particular Ehrlich is harshly criticized by them for not disclosing his data and bad workmanship. Assessing the literature of the past decades, their fundamental critique seems to be grossly exaggerated, but it is obvious that there is such a vast arsenal of studies with different results that everyone can pick out the results he likes best.

However, even faced with the different approaches described in this chapter, it seems not to be impossible to incorporate them in one all-embracing framework of individual decision making. Independent of the theoretical background and favor of a researcher, we think that the question of deterrence is empirically testable and that the existing studies should contain enough viable information to be a good basis for a deeper analysis. The numerous discrepancies, approaches, techniques, cultural differences, etc. literally demand to be exploited by a meta analysis - to reveal the reasons for this heterogeneity, how properties of the authors, their techniques or the studied populations affect the results; and if there is any basic deterrent effect when other influences are removed.

3 Meta Analysis

In questions of science the authority of a thousand is not worth the humble reasoning of a single individual.

Galileo Galilei

Do not believe in anything simply because you have heard it. Do not believe in anything simply because it is spoken and rumored by many. Do not believe in anything simply because it is found written in your religious books. Do not believe in anything merely on the authority of your teachers and elders. Do not believe in traditions because they have been handed down for many generations. But after observation and analysis, when you find that anything agrees with reason and is conducive to the good and benefit of one and all, then accept it and live up to it.

Buddha, Siddhartha Gautama

"Meta-analysis is used to provide a quantitative summary of the literature" (Rose, 2004). Or to be more precise: the primary goal of a meta analysis is to extract more information about a common question from a set of studies than the sum of each single study would yield. There are two ways to achieve this. The first is to make a new study using the accumulated raw data of all available studies while the second way is to exploit the reported results of these studies. When dealing with empirical studies about deterrence it is obvious that there is no way to use the former method of analysis because the data sources, on which these studies are based, are in most cases unavailable, not compatible, very heterogenous and unlikely to be summable at all in practice.

Nowadays, the most common form of meta analysis, at least in economics, is to statistically analyze the outcomes - which are interpreted as observations - and characteristics of empirical studies about a common topic. When narrowed down to the outcomes only, a meta analysis requires very similar studies which are most likely merely replications of each other. This kind of analysis was very popular and practical in the beginning[1]. Refer to Hedges and Olkin (1985) for more information about pitfalls, statistical methods and applications of these kind of meta analysis as well as Johnson et al. (1995). However, studies which share most of the non-sample properties are not common in sociology and economics. We explicitly include a very wide scope of studies which, as described in chapter 2, show very heterogenous characteristics and we control for most of these differences in the analysis. In general, we refer to the concept of meta regression analysis (Stanley

[1]The medical sector was the first to broadly make use of meta analyses because medical studies have, more or less, the same test design.

and Jarell, 2005) and regress the results of the included studies on study characteristics and other factors. This kind of meta analysis has become quite popular in economics; see Knell and Stix (2005); Longhi et al. (2005); Smith and Huang (1995); Stanley (2005b); Waldorf and Byun (2005) or Weichselbaumer and Winter-Ebmer (2005), for applications and further references.

The basic procedure can be divided into three steps. First, we describe the identification and collection of the relevant literature in section 3.1. Second, in section 3.2 we show how the relevant information is extracted from the collected literature and processed to make it accessible for further analysis. And finally, the meta analysis is performed in sections 3.5 and 3.6, preceded by a description of the data in section 3.3. Publication bias is considered in section 3.4. The project was interdisciplinary and worked on by two different teams, one in Heidelberg and the other in Darmstadt. The team in Heidelberg was responsible for the sociological and criminological literature, while the economic and miscellaneous literature was researched and processed by us.

At this point we should mention Pratt (2004) again, who also performs a meta analysis of empirical studies about crime. However, he uses a rather small set of studies and includes deterrence only as one among many other crime theories (a selection from theories can be found in subsection 2.1.3) and does focus on the social aspects of these crime theories. Furthermore, the set of covariates he uses is, compared with our data, quite small and his regressions are applied separately for each theory.

Entorf and Antony (2002) and Müller (2003) are preceding studies made in Darmstadt to demonstrate the feasibility and usability of a meta regression analysis of empirical deterrence. Although these authors use only very small sets of studies, they show that our task, to perform an extensive meta analysis, can be expected to be worthwhile.

3.1 The Search Process

. . . all ideas need to be heard, because each idea contains one aspect of the truth. By examining that aspect, we add to our own idea of the truth. Even ideas that have no truth in them whatsoever are useful because by disproving them, we add support to our own ideas.

John S. Mill, On Liberty, 1859

Nijkamp and Poot (2002) summarize the basic principle of the search process: the first step before any meta analysis can be carried out is the selection of the included studies. Coverage, defined as the extent to which the studies are representative of the targeted population and precision - the quality and the proximity to the topic at hand - are two very important issues. Unfortunately, these tend to vary inversely.

We explicitly and intentionally do cot comply with precision. Instead of selecting only studies with (supposedly) good quality, we collected all studies and incorporate quality issues into the meta analysis. Furthermore, deterrence theory in general covers so many different fields that a concise and narrow selection seems to be futile. While some authors of meta analyses search only in very few data bases[2] we include as many sources as possible. We do not only consider published studies and books but also searched, like van der Sluis et al. (2003), for unpublished studies (i.e., discussion or conference papers, reports, etc.).

The research took place in 16 criminological, sociological and economic data bases with the search terms "Abschreckung", "Generalprävention" and "deterrence" Depending on the data base, the search took place in one or more of the following categories: title, key words, abstract or full text. Additionally, we retrieved all articles which include references to Becker (1968) or Ehrlich (1973) (if the data base allowed this kind of search). Finally, each bibliography of every retrieved and relevant study was scrutinized for additional deterrence studies[3] (similar to Oosterbeek et al. (2004)). The search process resulted in 9422 references which were stored in a data base of references.

3.1.1 First Stage of Filtering

He that leaveth nothing to chance will do few things ill, but he will do very few things.

George Savile, 17th century

After eliminating multiple and obviously unsuitable references[4] in the first stage of processing, 3598 references remained, as depicted in table 3.1.

The number of references after the first stage does not say anything about the kind or quality of a data base because the elimination of duplicates followed no specific ordering. The fact that at least 787 references to relevant literature are taken from the bibliographies of acquired studies shows that, as depicted in subsection A.1.2, at least the coverage of older studies tends to be quite bad in these data bases.

3.1.2 Second Stage of Filtering

Nothing will ever be attempted, if all possible objections must first be overcome.

Samuel Johnson, Rasselas, 1759

In the second stage, the remaining references were categorized, distinguishing economic, sociologic and other references, and assigning each reference a number for its (presumed) empirical relevance[5]; from certainly being an empirical test of the deterrence hypothesis to unlikely to contain any relevant empirical information.

---

[2] Weichselbaumer and Winter-Ebmer (2005) and Knell and Stix (2005), for example, only searched in EconLit and some selected leading journals.

[3] Refer to subsection A.1.2 for further information about the impact of this additional step.

[4] These include, for example, literature about nuclear deterrence or market entry deterrence, anti-trust literature, book reviews, oviposition deterrence and similar biological subjects.

[5] The relevance was judged by the available information, like title, abstract, authors, journal, etc.

Table 3.1: Data bases in the research process before and after the 1st stage of processing

Data base	description	before	after
KrimDok	bibliography of criminological literature		
Sociological Abstracts	sociological abstracts		
Social Services Abstracts	sociological abstracts		
PsycINFO	psychological abstracts		
	lists of monographs		
		3568	281
EconLit	economic bibliography	850	205
ISI	Science Citation Index Expanded (SCI-EXPANDED)	1852	1190
NBER	National Bureau of Economic Research	32	3
RePEc	economic research papers	519	74
SSRN	Social Science Research Network	334	142
WISO-net	data base for economics and sociology	377	58
Ingenta	commercial content provider for scientific journals	1389	271
CiteSeer	Scientific Literature Digital Library	152	0
WoPEC	economic research papers	162	0
IZA	Institute for the Study of Labor	12	3
PsychARTICLES	psychological articles	19	0
IBSS	International Bibliography of Social Sciences	18	9
	other meta analyses and bibliographies		459
	bibliographies of relevant literature		787
	others		116
		9422	3598

The first five rows do not contain any detailed numbers because they could not be provided by the team in Heidelberg.
The last three rows were not part of the research in the data bases but were acquired during the whole project.

Like Beyleveld (1980) does, we exclude certain topics (besides the obviously irrelevant literature). We exclude studies dealing with the following topics.

- Specific deterrence. Articles which explicitly study the effect of punishment on recidivism of individual persons or groups. Some exceptions are made, for example, when the used data consisted only of prisoners but the focus was more on general deterrence.

- Psychological research. Research on the effects of punishment from a behavioral point of view. Also, no articles studying animals were included.

- Theoretical Studies. Articles which study methods, theories and other theoretical concerns regarding empirical deterrence without a relevant application.

- Studies solely about norm- or law abiding people (e.g., Orviska and Hudson (2003), immoral but legal behavior (e.g., Konar and Cohen (2000)) or implicit and abstract deterrence (e.g., the firm size in Alexander and Cohen (1999)).

We explicitly also include studies dealing with driving under the influence, speeding, cheating in classrooms, tax evasion, environmental offenses and violating requirements (safety measures, pollution limits, etc.). All[6] remaining studies were acquired to the best possible extend, following their presumed relevance. When a study was unavailable, we acquired more information from third parties in the internet to determine its relevance. Table 3.2 contains the result of the second stage process.

After acquiring[7] and looking through 1966 studies and carefully discarding 1632 other references to studies which were not attained, 840 relevant empirical studies remained. For further interesting statistics of these references see subsection A.1.1.

3.1.3 Main Resources

1966 studies have been acquired. Several studies are freely available (mostly newer working papers, government reports). Most studies were retrieved from electronic libraries of German universities (which had a subscription to the journal or had access to JSTOR or similar services). If electronic access was not available, they were scanned and e-mailed (postal deliverance for books) to us by a library of a German university, which had the journal or book in their shelves, for a small fee of about 5€ (8€ for books). This is a service nearly all German universities provide and is called SUBITO. 609 studies have been attained using SUBITO. Last but not least, some studies were attained from acquainted researchers (e.g., Mark Cohen, Horst Entorf and Hannes Spengler provided several studies) or we, if their study was unavailable to us, asked researchers per email. However, only very few answered and send us a copy.

[6] Although we could have done so, 203 references were not acquired because it would have been too expensive and any relevance was unlikely.

[7] For some statistics about the acquisition of the literature refer to tables 3.2 and 3.4.

Table 3.2: Category, relevance and availability in the 2nd stage of research process

Acquired category	relevance						
	before			after			
	strong	medium	low	strong	medium	none	sum
Economics	216	237	269	**268**	138	316	722
Sociology	171	316	232	**355**	135	229	719
Others	237	181	107	**217**	143	165	525
Sum	624	734	608	**840**	416	709	1966
Not acquired category	relevance						
	before			after			
	strong	medium	low	strong	medium	none	sum
Economics	127	58	152	0	0	337	337
Sociology	160	414	405	0	0	979	979
Others	162	39	115	0	0	316	316
Sum	449	511	672	0	0	1632	1632

The left columns contain the number before reading the acquired studies and getting more information about unattainable studies. The right columns describe the final categorization of the studies thereafter. Medium relevance means that the study is empirical and about crime but does not contain any relevant and usable deterrence variables. Low and no relevance indicate that a study is not empirical or not about crime. The bold cells give the number the relevant studies for the meta analysis.
The numbers for sociological studies of medium relevance is somewhat larger because studies, which were attained in Heidelberg, received no previous ranking and were all treated as being of medium relevance. Additionally, several studies from Heidelberg, which were not attained, could not be unambiguously classified as “not acquired” or “not relevant” and were assigned to the latter category.

After the acquisition of all studies and their analysis was finished, a new copyright law was passed in Germany. One of its consequences is that, if the copyright holder provides an (usually expensive) access to a study, the SUBITO service we used becomes prohibited. Since the average price for such studies is somewhere between 25$ and 45$, we would have been unable to reach such a level of coverage. Therefore, we are lucky that the passing of that law occurred, by chance, just after we finished this work.

3.2 The Data

The trouble with doing something right the first time is that nobody appreciates how difficult it was.

Walt West

The information the relevant studies provide has to be stored in a common data base to make it usable in a meta analysis. It is of utmost importance to have access to as much relevant information about the studies as possible. Therefore, we stored all relevant information in a unified, single data base. Since the teams worked independently in separated locations, the interface to enter the data, and the meta-data base itself, has to fulfill several conditions.

The Interface of the Data Base

The data base had to be accessed from anywhere at anytime on any platform but only from authorized persons. Several individual levels of rights had to be assigned to all people involved (from a guest-access which only allows the inspection of the data to full administration rights). Since new variables or data structure were introduced or modified during the course of the project, the data had to be easily and transparently maintained. Backup-management and a data format which is easy to handle was also important. Furthermore, functions to prevent users from making incorrect or invalid entries, calculating values and to guide users efficiently through the entry forms were implemented.

Therefore, the interface was written in HTML, CSS, JavaScript and PHP; the data base was realized with MySQL and both were hosted on a web site. MySQLDumper was used for the backup management and PHPMyAdmin for administering the MySQL data base. The web-interface consists of 6573 lines of code and provides an universal and flexible way to fit the individual needs of this project.

The Data Structure

The sets of variables to capture any relevant information was developed in Heidelberg and completed in Darmstadt. The data we collected can be categorized by two information criteria: study- and estimate-related data. The first part covers all general information about the study itself (characteristics of the publication, the author, the kind of study, the utilized data, quality aspects of the study, etc.) while the second part captures characteristics of each estimate[8] (characteristics of the independent and dependent variables, the used explanatory variables, aspects of the model, detailed information of the result, etc.). To evaluate the data we merged both parts by duplicating the study variables for each estimate. A study which provides n recorded results is thus represented by n rows (i.e., observations) in the date base - the first part of each row is exactly the same, the other part may be more or less different (depending on the results) as depicted in table 3.3.

Preparation of the Data

During the whole process it became necessary to manipulate, convert and to study the structure of the data, to recalculate certain values and to find and remove errors derived from inconsistent data. To make this possible we programmed a tool which provides the necessary features[9]. The most important are:

[8] By *estimate* we mean a result a study reports. For example, if a study tests the deterrence hypotheses for each of the seven Index I crimes in the United States with one regression, it contributes seven estimates to our data base. When we use the term *observation*, we usually refer to elements of our data base.

[9] To a certain extend, some features can be realized with the MySQL syntax as well.

Table 3.3: Excerpt from the data base

Variables	data				
Publication type	journal	journal	journal	…	journal
Author	Steven D. Levitt	Steven D. Levitt	Steven D. Levitt	…	Steven D. Levitt
Journal	Economic Inquiry	Economic Inquiry	Economic Inquiry	…	Economic Inquiry
Publication year	1998	1998	1998	…	1998
Author country	USA	USA	USA	…	USA
⋮	⋮	⋮	⋮	⋱	⋮
Time span of data (months)	264	264	264	…	264
Data base	UCR	UCR	UCR	…	UCR
Major problems (by reader)	no	no	no	…	no
Number of results	84	84	84	…	84
Independent variable	arrest rate	arrest rate	arrest rate	…	arrest rate
Dependent variable	crimes/pop.	crimes/pop.	crimes/pop.	…	crimes/pop.
Crime	murder, homicide	larceny	assault	…	auto theft
Number of explanatory variables	8	8	9	…	9
Error correction in model	yes	yes	no	…	no
Method used	OLS	OLS	OLS	…	OLS
⋮	⋮	⋮	⋮	⋱	⋮
Coefficient	-0.03	-0.154	-0.365	…	-0.457
Sd of coefficient	0.033	0.021	0.122	…	0.225
T-value				…	
Number of observations	819	819	963	…	963

The first column represents an (arbitrarily chosen) sample of the available variables in our data set. The other columns represent estimates from our data (thus, one column in the table resembles one row in our data base) and were drawn randomly from one study. The upper block contains the general information about the study and the lower block covers information about the individual estimates reported by the study.

- Converting the MySQL data bases into a flexible text-format and merging the two data sets into one.
- Calculating weights and, if possible, missing values (values of significance, converting statistics, adjusting signs, etc.).
- Detecting and deleting unused variables, generating new variables from others, converting the format of variables and conditionally renaming variables.
- Support the researcher in detecting inconsistent data (e.g., different or missing discipline of an author) and to provide an automated and menu-guided usage.

The tool was programmed in Java to keep the independency on any Operating System and consists of 2331 lines of code.

3.2.1 Data Entry Description

We are usually convinced more easily by reasons we have found ourselves than by those which have occurred to others.

Blaise Pascal, Pensée, 1657

In the following, we give some rough description of the data, some statistics about the data entry process and important information about the relevant variables we use in the meta analysis.

Time constrains prevented us to enter all 840 studies into the data base. Eventually, 700 studies were recorded, 350 by the team in Heidelberg and 350 by us. The team in Heidelberg recorded all estimates a study provided while we considered only one estimate for each crime and data set a study used (similar to van der Sluis et al. (2003)). This was necessary because we recorded all of the economic studies which often repeat an analysis with little variation in the variables to verify the robustness of a result[10]. Recording all these estimates would have taken too much time, and we decided that it would be better to include more studies (with a reduced number of estimates) than to use a reduced number of studies with all estimates[11]. In Rupp (2006) - a paper very similar to Bijmolt and Pieters (2001) - we showed that, under certain assumptions, taking random estimates leads to better results than taking the median (as Rose (2004) did) or mean value. Even more important, it would have been very difficult to record a mean or median value, when results also differ in other properties (e.g., are calculated with different subsets or specifications, which is usually the case).

Beside the results (in the literature often referred to as *effect sizes*), we also recorded as many properties as possible about the study design and the implemented methods. Although this is

[10] In economics and other fields, a result is usually assumed to be fragile if it can be reversed or mitigated by minor changes in the specification.

[11] "All explicitly reported estimates" would be a better expression because some studies refer to further unreported results. Usually these, in some cases up to several thousands, are not published for reasons of parsimony, or because they are all similar.

often done in meta regression analysis (e.g., Murphy et al. (2003); van der Sluis et al. (2003); Oosterbeek et al. (2004); Weichselbaumer and Winter-Ebmer (2005), etc.), the scope of recorded information in this meta analysis is unique. Results which are explicitly reprints from another study are not considered. Out of 7822 recorded values (including 1680 values favored by the author[12], but not randomly chosen), 7641 provide the sign (e.g., whether the reported estimate supports the deterrence hypothesis or not). To compare the reported values in a meta analysis, it is common to use the associated t-values reported by the studies (Stanley and Jarell, 2005). However, it would also be possible to use ordered logit (or probit) methods with the p-values of the results, like in Waldorf and Byun (2005).

For the 350 economical and "other" studies entered in Darmstadt, we have detailed statistics at our disposal. These statistics which provide information about the data entry process can be found in table 3.4. They can also be used to estimate the additional time required to enter the 140 omitted studies.

Table 3.4: Statistics of the data entry process of all 350 studies in Darmstadt

Statistic	sum	mean	sd	min	max
Pages	7090	20.26	15.82	3	121
Estimates	10695	30.56	66.05	0	764
Stored estimates	3140	8.97	12.29	0	127
Time to read	22605	64.59	36.71	5	245
Time to enter	10075	28.79	23.35	5	225
Time (total)	32680	93.37	50.75	15	420
Time per page		3.71	1.55	0.95	13.93
Time per estimate		5.48	3.82	0	25
Time saved	14345	40.99	105.29	-7.72	1289.06

All times are given in minutes.
Stored estimates: the estimates stored in the data base (one estimate per crime and source); *Time (total)*: total time for reading the study and entering the estimates into the data base; *Time per page*: time to read a study per page; *Time per estimate*: the time required to enter the data per recorded estimate; *Time saved*: the estimated saved time by the recording scheme (only one random estimate per crime and source) estimated by OLS (regressing Estimates on Time (total)).

3.2.2 The (Normalized) t-Value

Everything which is merely probable is probably wrong.

René Descartes

We chose to extend the usual procedure to rely solely on the reported, original t-values, by using all given values of significance (transforming them into t-values) and then normalizing them.

[12]Estimates with the property "favored by the author" were always recorded and it is noted in the data base whether they are also randomly chosen. Favored estimates which are not randomly chosen may be analyzed in the future and are neglected in this study. Fox example, Rose (2004) analyzes the results favored by the author of each study.

Normalization removes any systematic differences caused by the various t-distributions (which depend on the degrees of freedom and the implemented estimator), so that weighting them by the degrees of freedom (or just the sample size like in Knell and Stix (2005)), is not necessary. The procedure follows these rules:

1. If a t-value is reported, we take it as it is.
2. If a coefficient and its standard deviation is reported, we calculate the corresponding t-value, regardless of the used estimator[13]
3. If only the significance of a F- or χ^2- test is given,
   a) the value is transformed into a t-value, if the degrees of freedom are given;
   b) if the degrees of freedom are not given, they are approximated by the number of observations and covariates;
   c) if the number of observations and covariates is not given, it is approximated by a median number of 232 and 15 respectively.
4. If only the category of the p-value (not significant, 0.1, 0.05, 0.01 and 0.001) is given, the corresponding t-value is approximated by the following rules:
   a) an uniformly distributed number between the upper and lower limit of the category is chosen, representing the "exact" p-value[14];
   b) the corresponding t-value of this p-value is calculated, assuming two-sided tests[15], according to the following rules:
      i. using the degrees of freedoms, if reported;
      ii. if the degrees of freedom are not reported, they are approximated by the number of observations and covariates;
      iii. if these are not reported too, they are approximated by the median number of 232 and 15 respectively.
5. If an estimate supports the deterrence hypothesis it's t-value is provided with a negative sign and with a positive sign otherwise.
6. All t-values are normalized:
   a) a t-value is transformed into the corresponding p-value using the reported or approximated degrees of freedom and the t-distribution;

[13] We acknowledge that some of these values are only asymptotically t-distributed or even not at all. However, we prefer this inaccuracy, which should be quite small, to losing such estimates.

[14] However, this implies that even, in absolute values, the largest t-values calculated in this fashion are much smaller than the largest t-values reported by several studies; see table 3.5.

[15] If the study reports one-sided tests, this is considered in the data base accordingly.

b) this p-value is transformed by the inverse normal distribution to a value of significance which is independent of the number of degrees of freedom.

Due to limited precision, we were not able to normalize t-values[16] below -38. This affects 11 t-values, including eight which are favored by the author but are not randomly chosen; so, practically, only three t-values are affected. Some authors remove outliers in their meta analysis (e.g., Murphy et al. (2003) or Knell and Stix (2005)), but we want to keep them since we have no prior knowledge (except the sample size) of what could cause such outliers. However, since the relative difference between the t-values and their normalized counterparts can be quite large, we did not want to include them unadjusted and chose to transform these few outliers by the following formula, to conserve most of their relationship:

$$t_{\text{new}} := \frac{\log(|t_{\text{old}}|)}{\log(|t_{\text{min}}|)} t_{\text{min}},$$

where t_{min} is the smallest normalized t-value. This reduces the influence of these extreme values and retains the relationship between those values, at least at a logarithmic scale. In the case of the three values (excluding the eight favored values), this means the following transformations: $-582 \rightarrow -64.81134, -86.517 \rightarrow -45.40674$ and $-40.58823 \rightarrow -37.7018$. Some of the effects of this normalization procedure can be seen in tables 3.5 and 3.6.

Table 3.5: Comparison of the original and transformed (normalized) t-values

t-values	mean	median	min	max	%	#e	#s
Overall	−1.40	−1.37	−64.81	19.05	41.66	6530	663
Original	−1.66	−1.69	−64.81	15.16	44.93	2662	285
Calculated	−2.18	−1.47	−37.70	19.05	42.90	888	98
Transformed	−0.95	−0.89	−4.97	3.96	38.36	2980	328

Overall: all t-values; *Original*: all t-values reported in a study; *Calculated*: all t-values which were calculated by a given coefficient and sd; *Transformed*: p-values transformed into t-values.
% is the percentage of estimates which are consistent with the deterrence hypothesis and significant at a 5% level in a two sided test. *#e* is the weighted number of all valid estimates. *#s* is the number of studies the estimates are based on.

Weighting

As mentioned before, it was necessary to restrict ourselves to one estimate per crime and source per study. This makes it necessary to weigh the estimates in our data base in some way. In principle, there are three different approaches from which we chose the last one:

[16] In fact, this depends on the t-value and the degrees of freedoms simultaneously. Although being a very subjective limit, defining these t-values as outliers seems very practical.

1. Leave everything unchanged: i.e., use the unweighted estimates. However, studies which present numerous estimates would squeeze out the effects of studies with only a few (Stanley, 2005a). Moreover, "our" studies (i.e., those recorded in Darmstadt) would be underrepresented (see table 3.6).

2. Treat each estimate equally: weight each estimate in such a way that the sum of all weighted estimates of each study is equal to the total amount of results it contains. This would be an approximation of the case in which we record all results and would bias the analysis in favor of those studies with many results.

3. Treat each study equally: weight every estimate by the inverse number of the estimates in the data base belonging to the corresponding study. If a study recorded by the team in Heidelberg provides n estimates, it is weighted by $1/n$. A study recorded by us, of which m out of n estimates are in our data base, each is weighted by $1/m$. Therefore the sum of all weights of each study amounts to one.

Since "our" studies seem to differ significantly from the others, which can be readily appreciated by examining table 3.6, and the number of results per study varies substantially (from one to several hundred), we decided to use the latter weighting scheme.

Table 3.6: Weighted (normalized) t-values distinguished by the source of data

Source	obs.	%	mean	median	sd	min	max
Both, unweighted, not normalized	6530	100.00	−1.30	−0.91	7.77	−582	20.93
Both, unweighted	6530	100.00	−1.15	−0.91	2.70	−64.81	19.05
Darmstadt, unweighted	2320	35.53	−1.57	−1.20	3.61	−64.81	19.05
Heidelberg, unweighted	4210	64.47	−0.92	−0.75	1.99	−17.91	11.72
Darmstadt, weighted	2320	48.16	−1.76	−1.61	3.66	−64.81	19.05
Heidelberg, weighted	4210	51.84	−1.07	−1.11	2.36	−17.91	11.72
Both, weighted, not normalized	6530	100.00	−1.51	−1.37	4.51	−582	20.93
Both, weighted	6530	100.00	−1.40	−1.37	3.08	−64.81	19.05

The rows of the unweighted data refer to the first weighting scheme: leave everything unchanged. The rows of the weighted data refer to the third weighting scheme, which weights each study equally. Naturally, the number of observations and the extreme values are not affected by weighting. The % column indicates the fraction of the whole data set belonging to this row (measured either by the number of observations or sum of weights).

We acknowledge that there might be more weighting schemes possible, like using some impact factor of the publications (van der Sluis et al., 2003), removing heteroscedasticity (Murphy et al., 2003), using the sample size or the time frame (Knell and Stix, 2005) or adjusting for significance (Waldorf and Byun, 2005) or others, like the inverse variance of the results, number of results, number of regressors, R^2, etc. (Weichselbaumer and Winter-Ebmer, 2005). However, we refrain from using any of them because we do not want to mix different weighting schemes.

3.2.3 Adjustment of Variables

Nothing exists except atoms and empty space; everything else is opinion.

Democrites of Abdera, Diogenes Laertius IX

Our meta-data base contains many variables with missing entries. There are two main reasons for this: either the information is not available from a study (e.g., whether the used data is representative, some characteristics of surveyed people, the year the data was gathered, etc.) or the information is not applicable for a particular study (e.g., survey characteristics for a time series study, name of a journal for a book, etc.). Since we want to include as much data as possible in a multivariate analysis, we treat missing or not applicable information as zero values. Excluding these estimates would either result in an empty data set (there are always variables not applicable to a study) or restricting the analysis to very narrow subsets (with rarely more than a few dozens observations). Imputing variables is only reasonable for specific subsets and would be very difficult, even in these subsets. Although there are variables which could be imputed every applicable method would be questionable (e.g., the nationality of the author or whether the used data is representative). Also, variables are not independent. For example, the used data set will be correlated with the nationality and the field of the authors. Furthermore, it is not easy to identify the correct neighbors to generate the imputed values. Thus, we do not exclude estimates with missing values at all but treat missing information as unique zero-values[17].

We also removed those variables which had (almost) no entries. While removing unused variables does not pose any problems, removing variables with almost no entries is somewhat questionable. Losing information is never desirable but we have many variables with almost no variation which would make a further analysis even more difficult - especially in regard to the data mining methods. Therefore, we excluded those variables with less than seven (15) entries regarding study (estimate) related information. This amounts to a minimum variation of 1% (0.2%).

3.3 Descriptive Statistics

Basic research is like shooting an arrow in the air and, where it lands, painting a target.

Homer Adkins, Nature, 1984

Before delving into the multivariate dependencies, it is important to get a feeling for the data. This can be done in a very convenient way by illustrating certain properties and relationships of the data. Interdependence of certain variables with the (normalized) t-values are also of interest. A full list of all available variables with summarizing statistics can be found in the codebook in section B.1.

[17] There are some rare cases when a variable can take the value zero (e.g., the percentage of males in a sample) but these are negligible.

3.3.1 Study-Related Description

Yet to calculate is not in itself to analyze.

Edgar A. Poe, The Murders in the Rue Morgue, 1841

Herein are contained all tables and graphs which improve the knowledge about study-related variables; i.e., variables which cannot vary within each study. We begin with a summary of the countries related to the studies contained in our data base.

Table 3.7 makes it obvious that the deterrence literature is dominated by authors, data and journals from the United States of America[18]. More than three quarters of all authors worked in the USA at the time of writing, used U.S. crime data and published in U.S. based journals - even 20.41% of all authors who did not work in the USA at the time of writing used U.S. crime data. This is not unexpected because there are many data sets (crime data and surveys) available for the USA; some starting in the early 1930s (although the reliability of those early data sets is questionable). The United Kingdom, Canada and Germany (and partially Australia) make up for the major part of the rest. Germany, and maybe the Netherlands, seems to be the only non-anglo-saxon country with a relevant portion of published studies. However, this may result from, at least to some extent, the inclusion of German expressions in our search terms.

Table 3.7: Most frequent countries

Country	workplace	%	studied	%	published	%
Australia	16	2.20	13	1.82	4	0.57
Canada	34	4.68	28	3.92	23	3.29
Finland	6	0.83	7	0.98	0	0.00
Germany	22	3.03	19	2.66	31	4.43
Israel	9	1.24	4	0.56	0	0.00
Netherlands	8	1.10	8	1.12	18	2.57
Other	35	4.82	47	6.58	14	2.00
Sweden	6	0.83	9	1.26	3	0.43
Switzerland	8	1.10	8	1.12	1	0.14
United Kingdom	30	4.13	33	4.62	56	8.00
USA	552	76.06	538	75.35	550	78.57

The columns may not sum up to 700, since some studies are written by authors from different countries, analyze different countries simultaneously or the information is not available; only the country of publication is unambiguously determined. The percentages are calculated from these data and always sum up to one; all deviations are based on rounding.

In fact, from 451 studies which use a public data base, 163 (36.14%) use the Uniform Crime Report (UCR) - the official data base about crime in the USA, compiled by the FBI. The other 63.86% are made up of surveys (e.g., the National Household surveys or National Longitudinal

[18] To determine the country of publication of a journal, we first referred to the country of the founding editor and, if that information was not available, of the leading editor, the editor who receives the manuscripts, the majority of co-editors or the publisher.

Survey of Youth for the United States), the German *Polizeiliche Kriminalstatistik* (PKS) and other data sets. All other either do not specify the origin of their data, use data collected by themselves (mostly surveys) or use confidential data.

Overall, 848 authors are involved in the 700 studies we have included in our meta-data base. Most of them appear only once (75.83%) or twice (14.03%). The authors who contribute at least six studies are given in table 3.8: 64 authors (7.55%) contribute three to five studies and are not shown. Merely 2.59% of all authors are involved in 22.71% of all studies. It is not surprising that William C. Bailey and Raymond Paternoster come first, both being sociologists and (Co-) authors of more than 15 studies; the economists Steven D. Levitt and Dale O. Cloninger following close behind. From these top 22 authors only the German Horst Entorf and Finnish Matti Virén did not work in the USA at the time of writing.

Table 3.8: Most frequent authors

Author	N	Author	N
Bailey, William C.	20	Piquero, Alex R.	8
Paternoster, Raymond	16	Rasmussen, David W.	8
Levitt, Steven D.	12	Waldo, Gordon P.	8
Cloninger, Dale O.	11	Hakim, Simon	7
Grasmick, Harold G.	10	Witte, Ann D.	7
Ross, Laurence H.	10	Entorf, Horst	6
Benson, Bruce L.	9	Erickson, Maynard L.	6
Nagin, Daniel S.	9	Marvell, Thomas B.	6
Chiricos, Theodore G.	8	Pogarsky, Greg	6
Ehrlich, Isaac	8	Tittle, Charles R.	6
Gibbs, Jack P.	8	Virén, Matti	6

All authors with at least 6 contributing studies are listed.

The disciplines of all authors are, more or less, evenly spread (43.23% economists, 40.55% sociologists, criminologists and jurists). Regarding the disciplines of the publisher[19] the relationship is reversed (36.34% and 45.35) which is not surprising, because economists more readily publish in journals outside of their subject than sociologists and criminologists do[20]. Table 3.9 depicts this in more detail.

All in all, the 700 studies we recorded are mostly published in journals (86.29%), followed by working- and conference papers (5.71%) and books (5.57%). 2.43% are not classified. 199 journals publish the 604 journal articles, whereas 120 journals contribute only a single article; the major part, namely 326 articles (53.97%), stem from just 25 journals (12.56%) and are shown in table 3.10. The Criminology-Journal stands out by contributing 30 studies, fifty percent more than

[19] This can be, for example, a journal, a department or an institution.

[20] Jurists are an exception here: 27% of their studies are published in economic media, which is the largest cross-discipline share of all. However, this discipline just appears in 26 studies in our meta-data base.

Table 3.9: Disciplines of authors and publishers

Publisher discipline	Author discipline						
	Law	Crim.	Econ.	Soc.	Psych.	Other	Sum
Law	7	7	15	5	1	2	37
Criminology	8	43	15	64	3	10	143
Economics	7	1	207	10	2	12	239
Sociology	2	15	22	80	2	13	134
Psychology	0	0	0	5	12	3	20
Other	2	7	29	11	8	44	101
Undefined	0	0	3	0	0	0	3
Sum	26	73	291	175	28	84	677

While all disciplines of the publisher are classified, some authors could not be categorized (the missing 23 studies). Discipline of the publisher is unique while multiple authors can contribute multiple disciplines to a study-entry.

the second most frequent journal. All major disciplines are present in the top five (except the law discipline, which first appears in the 13th place). The listed journals make up for 43.75% of all law articles, 76.19% of all criminological, 53.08% of all economic, 67.2% of all sociological and 28.89% of all other journal articles.

The oldest study in our data base was published 1952, the newest in 2006. Not surprisingly, the data used in the studies is older, since it requires some time to collect and analyze the data, newer data is not available or simply not of interest. Some studies are even especially interested in old data (e.g., those which include data from the 19th century). The distributions of the year of publication and the mean year[21] of the used data are shown in figure 3.1. The median time span, conditional on multiple points of time, is exactly 120 months.

A steady stream of empirical studies about deterrence started in the late sixties, reaching a first peak in the late seventies with the heated discussion fueled by Ehrlich (1973) and Ehrlich (1975a). With the exception of some years, the number of publications slowly receded until the mid nineties when it reached new heights and kept its pace until today.

The steady flow of studies can, at least partially, be viewed under two different aspects: a diminishing interest in the deterrence research and a larger acceptance of deterrent measures in other fields of research (like unemployment in Smith et al. (1992); Witt et al. (1998) or Levitt (2001)) to minimize an omitted variable bias. This can be verified by figure 3.2. The number of studies which use deterrence variables only as covariates can be interpreted as a lower bound, because we classified deterrence variables as covariates only in the case the authors do not interpret[22] the corresponding coefficients. Furthermore, studies which do not focus on deterrence but include deterrence measures as covariates had a lower probability to be detected during our research.

[21] The mean year is the mean of the first and last year the data refers to, which is available for 593 studies.

[22] For example, if the effect of unemployment on crime is the focus of the study but the authors spend a considerable amount of space on interpreting the negative effect of the arrest rate as a deterrent factor, the according observation

Table 3.10: Most frequent journals

Journal	n	category
Criminology	30	criminology
Journal of Law and Economics	20	economics
Law and Society Review	19	sociology
Applied Economics	17	economics
Accident Analysis and Prevention	17	other
Journal of Criminal Justice	16	criminology
Journal of Criminal Law and Criminology	16	criminology
American Economic Review	15	economics
American Sociological Review	15	sociology
Journal of Research in Crime and Delinquency	15	criminology
Social Forces	14	sociology
Social Science Quarterly	14	sociology
Journal of Legal Studies	14	law
American Journal of Economics and Sociology	12	sociology
Review of Economics and Statistics	12	economics
Social Problems	10	sociology
Economic Inquiry	10	economics
Journal of Quantitative Criminology	9	criminology
Journal of Studies on Alcohol	9	other
Southern Economic Journal	8	economics
International Review of Law and Economics	8	economics
Journal of Public Economics	8	economics
Crime and Delinquency	7	criminology
Journal of Political Economy	7	economics
Journal of Behavioral Economics	7	economics

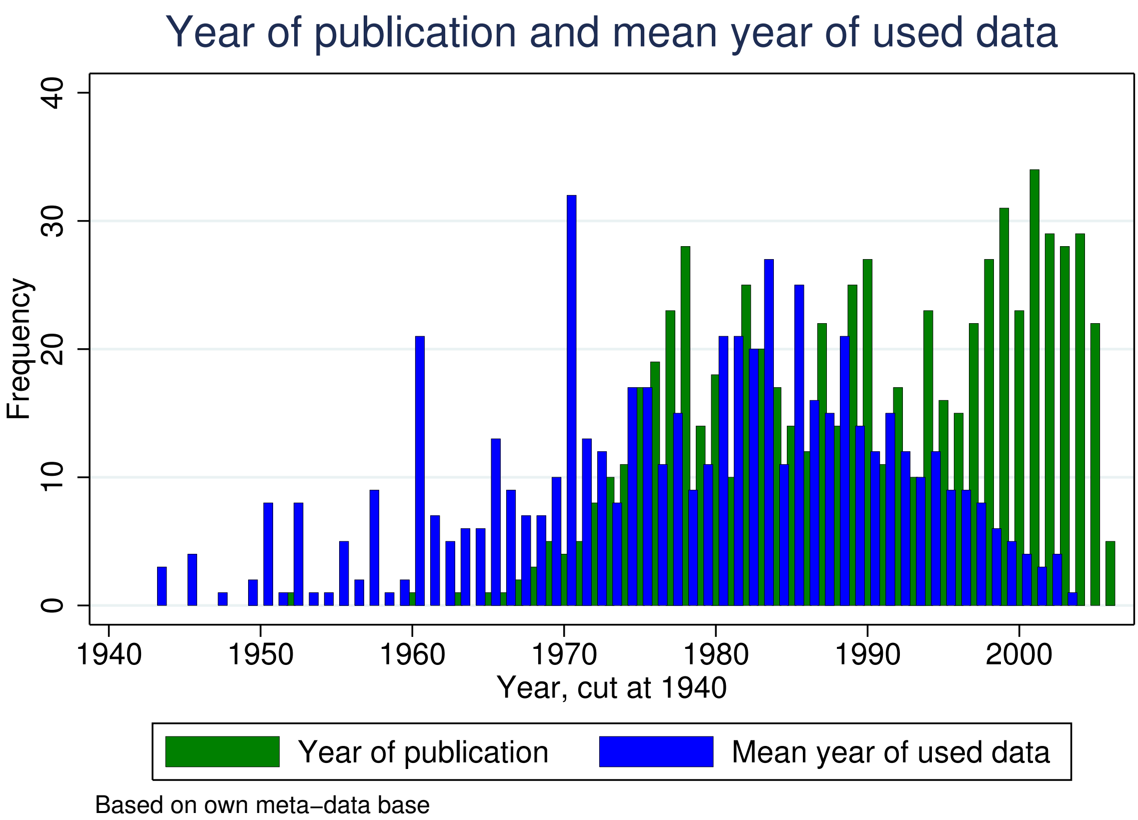
All journals with at least 7 contributing studies are listed.

The kind of data employed is spread rather evenly over the major categories - time series, cross sections, surveys and panel data (in that order) - while experiments appear only sparsely. Aggregated, almost two thirds (65.81%) is "typical" crime data, while 23.87% and 9.16% is survey and experimental data. Consequently, this is also reflected by the studied objects: counties, states, people and nations (in that order) are studied in most cases. 61.96% of all studies implement regional data and 30.49% of the data is derived from people and groups. These observations are visualized in figure 3.3.

We also included several variables to catch some quality related measures for each study. From some of these, we constructed a quality index starting at 0 (neither the author nor the reader report any problems) and ending at 8 (author and reader report major problems in the study). Summarizing these into three categories, we can say that 189 studies (27%) are rather flawless, 438 do contain some problems (62.57%) and 73 studies are problematic (10.43%). We stress that

in our meta-data base is not considered being only a covariate.

Figure 3.1: Year of publication and used data



this measure should not be overrated, because it is not easy to decide how many problems a study may have and how severe they are - it depends heavily on what the author does, reports and how frankly he is to the reader. Eventually, there is no such thing like a flawless empiric study without any problems.

Finally, we want to report that 79% of all studies use covariates and, therefore, control for other influences. In table 3.11, we show the key statistics of the number of results (i.e., the observations in our meta-data base) reported in the studies.

3.3.2 Estimate-Related Description

There are no facts, only interpretations.

Friedrich W. Nietzsche, Notebooks, 1900

In this subsection, all reported values are weighted values to compensate for the unequal recording scheme we had to use (refer to subsection 3.2.1). Consequently, the reported frequencies are all weighted and may reflect the true frequency only approximately (rounding error) and are not equal the number of observations in the meta-data base.

We included all relevant studies we could find, whether they studied deterrence or simply used

Figure 3.2: Year of publication diversified by focus of the studies

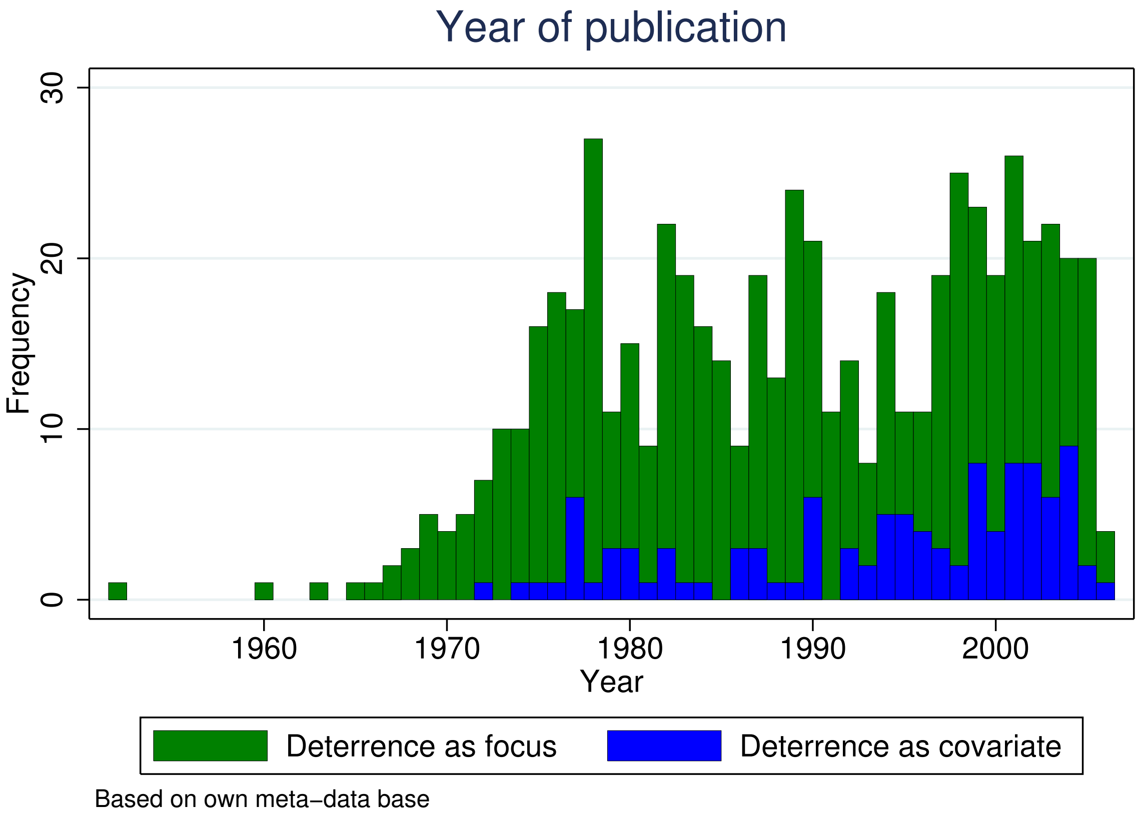


Table 3.11: Number of reported results in the studies

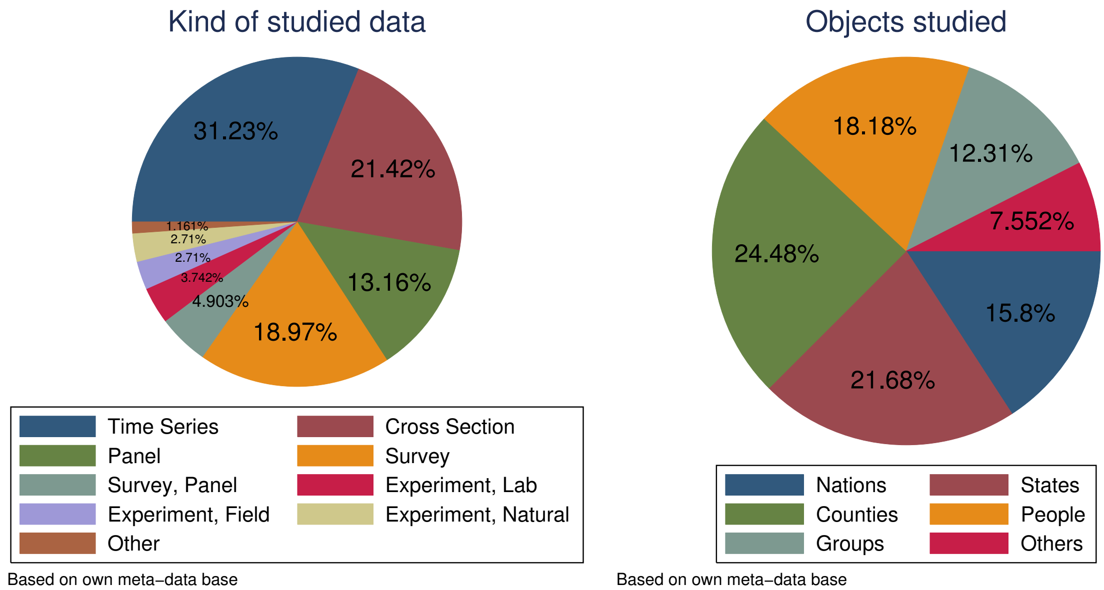
Method	mean	median	sd	min	max	n
Bivariate	4.161	0.0	14.014	0	155	700
Multivariate	17.682	6.0	47.692	0	764	700
Bivariate⋆	12.240	4.5	21.908	1	155	238
Multivariate⋆	21.716	8.0	52.024	1	764	570

Bi- and *Multivariate* refers to the kind of analysis the deterrence variable is subject to. Rows with ⋆ are conditional on the existence of such observations. Many authors report that they have calculated much more results (up to many thousands) but do not report them for reasons of parsimony or lack of available space. *n* refers to the number of corresponding studies.

such variables as covariates for other problems. The latter was the case in 14.23% of all observations. In most cases (54.92%), the observations relate to an analysis using crime data, in 24.56% they relate to a survey, 12.26% belong to an experiment or miscellaneous method[23]. Observations corresponding to deterrent effects of the death penalty are distinguished from the rest, although

[23]The numbers given in subsection 3.3.1 are based on study-properties; the small deviations stem from studies which report results based on different kind of data.

Figure 3.3: Kind of the employed data and studied objects



they would belong, in most cases, to the crime data category, and make up for the remaining 8.26%. In table 3.12, we report the most commonly used exogenous variables. Those which appear in less than 1% of all data are not listed and merged with the category "other"[24].

In most cases, the effect of the death penalty is measured by the execution rate (57.65%) or by uncategorized measures (20.87%), like the coverage of executions by the media (Forst, 1977; Stack, 1990) or the time interval between sentence and execution (Bailey, 1980a). In the case of crime data, variables which measure the probability of punishment dominate. After the "other" variable, the arrest rate (17.11%) comes first, followed by the number of police officers (13.86%), the clearance rate (9.33%), police expenditures (8.10%), the conviction rate (6.88%) and the two regime shift[25] variables (6.28% and 6.10%). Only the latter and the mean sentence length (7.43%) measure the severity of punishment. This is not surprising, since the severity of punishment is often not available or not applicable (for example, if the probability and severity of punishment are measured on different aggregation levels).

In surveys, the most prominent question is the expected probability of arrest by the police (28.26%), followed by the uncategorized variables (18.19%) and the expected probability of punishment by justice (17.58%). The severity of expected punishment is used more frequently (13.14%), as expected, than in the case of crime data. The same can be observed for experiments (and miscellaneous methods), where the actual variation of the detection probability (30.67%) is directly followed by the actual variation of the severity of punishment (18.23%). After the uncategorized variables (17.97%), the same order is observed for experimental variation of the

[24] We analyzed every "other" category and extracted every variable which appeared quite frequently. Therefore, there is no major variable hidden in that category.

[25] These are indicator variables which measure a shift in the deterrence regime; e.g., the introduction of new laws.

probability (15.76%) and severity of punishment (15.65%). This is quite reasonably, because it is much easier to ask people about the severity of punishment or to vary it in experiments than to attain such information in a reliable and appropriate fashion from existent crime data.

Table 3.12: Most frequently used deterrence variables diversified by the kind of data

Variable	n	%	studies
Death penalty	792	100.00	82
Execution rate	457	57.65	58
Other	165	20.87	22
Existence of death penalty	91	11.44	14
Percentage of all convictions	65	8.24	11
Conviction rate	26	3.25	4
Crime Data	3589	100.00	410
Other	629	17.52	122
Arrest rate	614	17.11	115
Police strength	497	13.86	88
Clearance rate	335	9.33	66
Police expenditures	291	8.10	51
Conviction rate	247	6.88	74
Probability dummy	225	6.28	32
Severity dummy	219	6.10	32
Mean sentence length (sentenced)	137	3.83	48
Mean sentence length (served)	129	3.60	35
Inspections	73	2.04	11
Incarcerations (absolute or per capita)	63	1.75	19
Incarceration per crime	56	1.55	13
Incarceration rate	45	1.24	20
Fine	44	1.23	18
Convicted per crime	42	1.16	11
Surveys	2534	100.00	175
Probability of detection by police	716	28.26	84
Other	461	18.19	46
Probability of punishment by justice	445	17.58	55
Severity of punishment	333	13.14	55
Probability of punishment by friends or family	157	6.20	25
Severity of punishment by friends or family	92	3.60	16
Probability of punishment by others	85	3.37	13
Previous experience with police or justice	82	3.22	6
Probability of punishment by employment law	59	2.33	9
Type of punishment	48	1.90	8
Severity of punishment by others	44	1.75	8
Severity of punishment by employment law	42	1.66	5
Probability of other kind of punishment	40	1.56	5
Probability of detection by others	39	1.54	7

continued on the next page...

. . . last page of table 3.12 continued

Variable	n	%	studies
Probability of detection by friends or family	31	1.22	5
Severity of other kind of punishment	30	1.17	5
Experiments	337	100.00	89
Actual variation of detection probability	103	30.67	30
Actual variation of punishment severity	61	18.23	20
Other	61	17.97	17
Experimental variation of punishment severity	53	15.76	19
Experimental variation of detection probability	53	15.65	19
Net gains (person is delinquent)	8	2.34	2

n is the number of occurrences and % the corresponding percentage (both weighted). *studies* is the number of studies which use that variable. If not listed separately, rates do also include the absolute numbers. Variables for surveys always relate to expectations. Some columns may not sum up exactly due to rounding or multiple entries. All entries with less than 1% (13 variables) were merged with the category *Other*.

end of the table 3.12

As with the deterrence variables used in the studies, we summarize the variables measuring crime in table 3.13. Since crime data is studied in most cases, the crime rate and the number of reported crimes are very prominent (first and fourth place with 44.92 and 11.21 percent). The second place is taken by the self reported delinquency rate (17.98%) which is usually often used in surveys. Accidents (4.25%) are commonly used in DUI-studies to measure the extent of drunk driving, while the violation of prescriptive limits (3.04%) belongs to studies which focus on environmental offenses (e.g., oil pollution, exceeding fishing quotas, etc.).

Table 3.13: Most frequently used endogenous crime variables

Variable	n	%	studies
All	7259	100.00	699
Crime rate	3261	44.92	319
Self reported delinquency rate	1305	17.98	132
Other	993	13.67	117
Reported crimes (absolute number)	814	11.21	85
Accidents	309	4.25	33
Probability of future delinquency (respondent is delinquent)	282	3.88	31
Violations of prescriptive limits	221	3.04	24
Probability of fictitious delinquency (respondent is delinquent)	79	1.09	10

n is the number of occurrences and % the corresponding percentage (both weighted). *studies* is the number of studies which use that variable. Some columns may not sum up exactly due to rounding or multiple entries. All entries with less than 1% (9 variables) were merged with the category *Other*.

We generally use the term *crime* although this is not literally correct. While crimes are studied in most cases (91.40%), misdemeanors (9.85%) and other offenses (e.g., deviant behavior, cheating in class rooms, etc.) are also studied, as shown in table 3.14. The classification is not always easy, because offenses may be judged differently across countries (and even between states). For example, we treated driving under the influence as crimes, because it is in most cases measured by fatal accidents with intoxicated drivers or offenses punishable by jail sentences[26].

Table 3.14: Formal severity of the studied offenses

Variable	n	%	studies
All	7260	100.00	699
Crimes	6636	91.40	646
Misdemeanors	715	9.85	80
Formal deviant behavior	198	2.73	26
Breaking of rules	162	2.23	16
Informal deviant behavior	118	1.62	19
Other	103	1.42	11

n is the number of occurrences and % the corresponding percentage (both weighted). *studies* is the number of studies which use that variable. Some columns may not sum up exactly due to rounding or multiple entries. Drunk driving is generally treated as a crime. All entries with less than 1% (0 variables) were merged with the category *Other*.

Besides these broad categories of the severity of the recorded crimes, we have, of course, recorded each studied crime in more detail. These are listed in table 3.15. It is not astonishing that the Index I crimes play a very dominant role - as mentioned before, most studies use U.S. crime data and especially the UCR - homicide, rape, assault, robbery, larceny, burglary, vehicle theft (and arson). Besides these, DUI (12.33%) and tax evasion (7.09%) are also studied fairly often. All other offenses are studied less frequently; many offenses appear only a few times (e.g., arson, cheating or airplane hijacking).

Authors implementing multivariate methods usually use covariates to control for other effects. We have recorded all those covariates and report these in table 3.16. This table does not include any other deterrence variables[27], although they are available in our meta-data base. The *Other*-category was scrutinized and new variables were generated from any covariates which appeared reasonably often. Income, unemployment, race and age are the most frequently used covariates. They are commonly available in various data bases for many countries and are common proxies for different incentives of crime. Income (40.80%) can be interpreted in two quite different ways: as a proxy of wealth, which increases the opportunity costs of crime, or as a proxy for the property value, which increases the incentives of crime (see subsection 2.4.2). Similar, unemployment (37.21%) can also be interpreted in two ways: either increasing crime (more available time, lower opportunity costs) or decreasing it (decreasing property values). Race (31.61%) is most often

[26] We are aware of the fact that driving with low but illegal BAC levels, is usually a misdemeanor.

[27] When regressing a crime rate on the arrest rate and sentence length, the latter is technically a covariate when we look at the significance of the arrest rate.

Table 3.15: Most frequently used endogenous crime categories

Variable	n	%	studies
All	7259	100.00	698
Homicide	1582	21.49	250
Robbery	919	12.65	215
Larceny (Index I, general)	902	12.43	200
Driving under the influence	895	12.33	95
Burglary	891	12.27	197
Assault	724	9.98	179
Other crimes	689	9.50	104
Vehicle theft	645	8.89	145
Tax evasion	515	7.09	54
Other	501	6.90	62
Rape	494	6.81	129
Overall crime	469	6.46	69
Drug (general) related	309	4.25	56
Other Misdemeanors	277	3.82	40
Fraud	269	3.70	48
Larceny (more than 50€ or 50$)	234	3.23	46
Petty theft	191	2.63	37
Environmental crimes, Violations of prescriptive limits	166	2.28	18
Negligent assault	138	1.91	39
Damage to property	98	1.14	19
Manslaughter	90	1.24	18
Drug (soft) possession	80	1.10	9
Speeding	74	1.02	9

n is the number of occurrences and % the corresponding percentage (both weighted). *studies* is the number of studies which use that variable. Some columns may not sum up exactly due to rounding or multiple entries. All entries with less than 1% (11 variables) were merged with the respective category *Other*.

used to differentiate black and white people in the USA. It is often assumed[28] that black people are more often involved in crimes than white people. Age (27.87%) and youth (20.33%) are also used very often since it is common knowledge that (detected) crime is decreasing with age after reaching its height for youths and young adults (see subsection 2.3.4). Fixed- and random effects are not usually called covariates. However, since these are used to pick up effects of unobserved heterogeneity by the introduction of dummy variables, they fit in this category quite well. Random effects seem to play no significant role in the deterrent literature. Since individual data only rarely studied, compared to aggregated data, is is not surprising that personal information like risk propensity, previous incarcerations or the social class appear at the bottom of the table. However, norm acceptance and morality appear relatively often.

[28] Gyimah-Brempong (1986) gives a good summary of what leads to this assumption and how it dissolves when controlling for other appropriate influences.

Table 3.16: Most frequently used covariates

Variable	n	%	studies
All	4579	100.00	543
Other	4110	89.76	501
Income	1951	40.80	223
Unemployment	1704	37.21	211
Race (black, white, etc.)	1448	31.61	179
Age	1276	27.87	163
Youths	931	20.33	119
Sex	895	19.54	127
Population (-growth)	749	16.37	98
Fixed Effects (time)	743	16.22	90
Fixed Effects (spatial)	614	13.40	81
Education	596	13.02	76
Income inequality (Gini)	541	11.81	68
Urbanity	527	11.52	73
Poverty, welfare	408	8.91	54
Marital status	375	8.18	51
Time trend	336	7.34	45
Nationality	197	4.30	26
Labor force	189	4.12	25
Personal characteristics	179	3.91	24
Norm acceptance	139	3.03	25
Consumption	126	2.75	16
Alcohol (-consumption)	121	2.64	17
Property value	113	2.48	15
Religion	107	2.33	17
Morality	98	2.15	15
Previous convictions	98	2.14	14
Social Integration	90	1.97	15
GDP	88	1.92	12
Miles (or km) driven	81	1.78	11
Drug (consumption)	78	1.70	11
Random Effects	61	1.33	10
Risk propensity	55	1.20	9
Previous incarceration	53	1.17	7
Social class	48	1.05	7

n is the number of occurrences and % the corresponding percentage (both weighted). *studies* is the number of studies which use that variable. Some columns may not sum up exactly due to rounding or multiple entries. All entries with less than 1% (8 variables) were merged with the category *Other*.

end of the table 3.16

The studies use a vast arsenal of statistical methods. From simple comparison of percentages to complex multivariate estimators. These methods are shown in table 3.17. The bivariate methods are dominated by the Pearson correlation (40.96%) and simple differences (typically before/after comparisons, 16.98%). Most of the multivariate methods are OLS (48.07%) or 2SLS (13.58%) regressions. Although similar to the bivariate case, it is not comparable because a very large class of estimators is contained in these regressions. OLS (2SLS and others as well) does only refer to the regression model, not to the specific model specification. Usage of error-corrections, lag structures, transformations and weights were recorded in other variables.

Table 3.17: Most frequently used statistical methods

Variable	n	%	studies
Bivariate	2451	100.00	242
Pearson correlation	1004	40.96	119
Differences (of means, percentages, etc.)	416	16.98	39
Gamma	208	8.49	21
ANOVA	184	7.50	15
Other	161	6.58	17
χ^2-test	147	6.00	12
OLS Regression	88	3.61	13
Kendall's τ	73	2.99	7
T-test (dependant)	48	1.97	5
T-test (independent)	46	1.89	5
Spearman's ρ	40	1.62	4
Point-biserial correlation	35	1.42	5
Multivariate	4808	100.00	571
OLS regression	2311	48.07	316
2SLS regression	653	13.58	96
Other	524	10.89	90
Logit or probit	461	9.59	58
ARIMA-models	278	5.79	32
Tobit regression	206	4.27	28
GLS regression	101	2.09	14
Path analysis	80	1.67	14
Poisson regression	69	1.43	11
Other ML-methods	65	1.36	9
VAR-models	60	1.25	9

n is the number of occurrences and % the corresponding percentage (both weighted). *studies* is the number of studies which use that variable. Some columns may not sum up exactly due to rounding or multiple entries. All entries with less than 1% (5 and 5 variables) were merged with the respective category *Other*.

3.3.3 Description of the Endogenous Variables

What men really want is not knowledge but certainty.

Betrand A. Russell

We have recorded four different properties which characterize the effect of the analyzed deterrence variables: the sign, the p-value and the (normalized) t-statistic[29] on the estimate-level and the opinion of the author on the study-level. As in subsection 3.3.2, the reported frequencies are all weighted and do not represent the exact numbers in the meta-data base. The sign is available for 7085 observations: 5241 (73.98%) observations have a negative (and thus tend to be in accordance with the deterrence hypothesis) and 1844 (26.02%) have a positive sign. However, not all observations have a usable test-statistic; only 6530 p- and t-values could be recorded or transformed. Their distribution is illustrated in table 3.18 and figure 3.4.

Table 3.18: Distribution of the p-value categories

p	−.001	−.01	−.05	−.1	−1	+1	+.1	+.05	+.01	+.001	All
N	672	835	1216	279	1820	1142	52	282	138	96	6530
%	10.28	12.79	18.62	4.27	27.87	17.48	0.8	4.32	2.11	1.46	100

The first row corresponds to the categories of the p-values (two-sided tests), diversified by the sign of the result. The second and third row contain the frequencies and percentages of the occurrences; due to weighting and rounding the columns may not sum up exactly.

While 4820 (73.81%) (normalized) t-values are negative (1710 or 26.19% are positive), 2720 (41.66%) are even smaller than −1.96, the typical 5%-level of significance (however, 506 or 7.75% are larger than 1.96). Although simple vote-counting[30] has become obsolete it is worthwhile to look at figure 3.4 which presents the mean and median, as well as the quantiles (table 3.19) of the (normalized) t-values. They indicate that it is more plausible to keep the deterrent hypothesis than do discard it. Assuming that there is no deterrent effect (i.e., under the null hypothesis), the probability to observe a value of −1.4 or lower is about 8% (p-value of one-sided test). However, assuming that the existence of a deterrent effect would lead to a "true" value of −1.96, than the probability of observing a value of −1.4 or larger would be about 29%[31]. Hence, the observed value is more likely to reject the null hypothesis of no deterrence than the null hypothesis of an deterrent effect which would lead to an average value of −1.96.

The obvious spikes at the usual levels of significance can be partially explained by the transformation procedure of the p-values into t-values (subsection 3.2.2). The dent between the two major spikes is located at the mean value and may be explained by a possible publication bias (as explained in section 3.4).

It is remarkable that the mean and median values are almost identical and the 10% and 90% quantiles lie symmetrically around the mean. The distance from the positive 99% quantile and maximum to the mean is much shorter than on the negative side. The distribution is heavy tailed on the negative side and short tailed on the positive.

[29] See subsection 3.2.2 for details about the normalization and imputation procedure.
[30] Vote-counting (i.e., how often a hypothesis is supported by the studies) is known to suffer strongly from error of the second kind, which increases with the number of studies (Hedges and Olkin, 1985).
[31] Both probabilities would be equal if the "true" t-value is −2.8.

Figure 3.4: Distribution of the (normalized) t-values

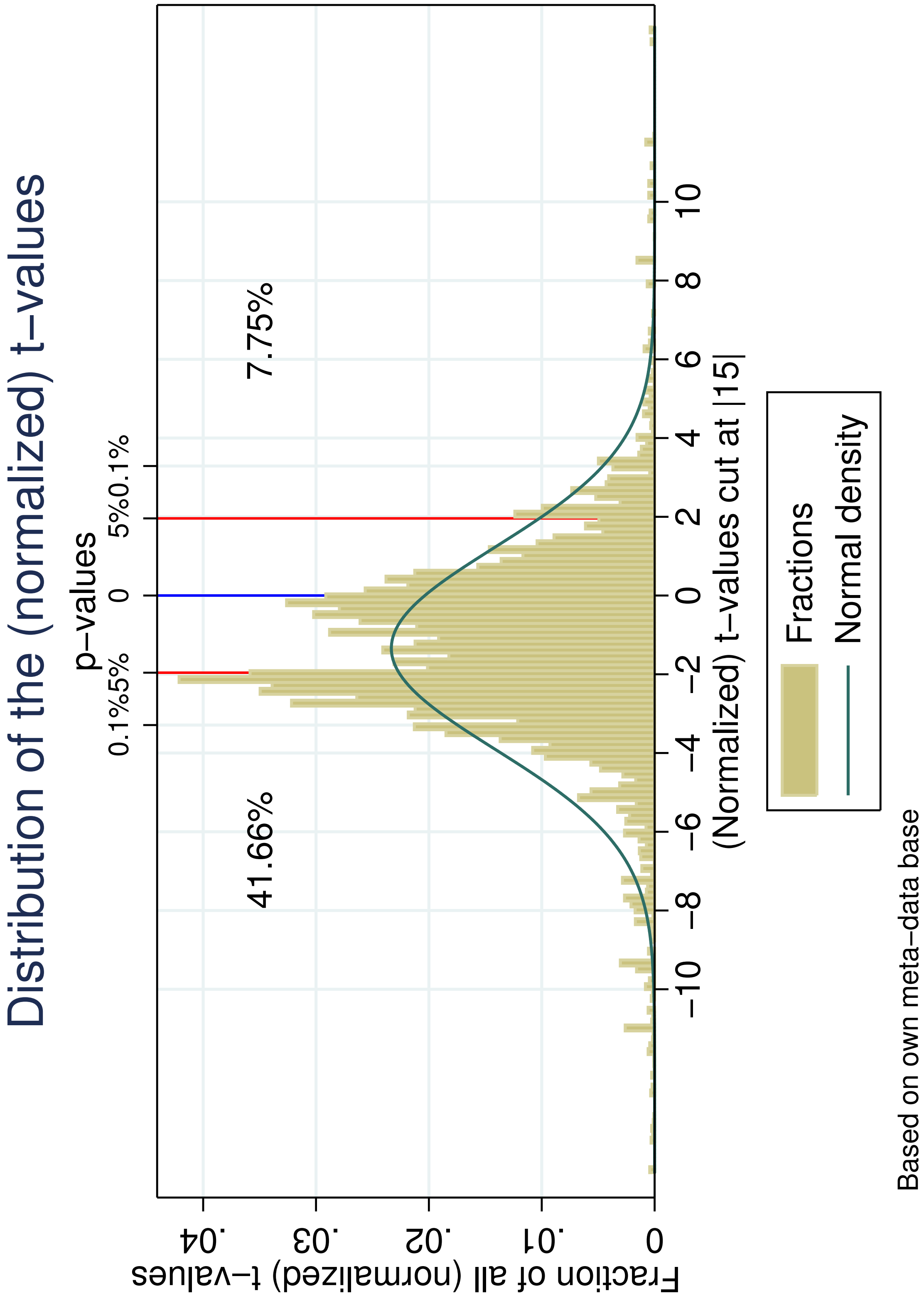


Table 3.19: Statistics of the (normalized) t-values

statistic	min	1%	10%	50%	90%	99%	max	mean	sd	n
value	-64.81	-10.92	-3.89	-1.37	1.37	5.85	19.05	-1.40	3.07	6530

Beside these objective criteria, we have also the opinion of the authors[32] at our disposal. We recorded the opinion for three broad crime categories (violent-, property- and other crimes) with 5 possible values - from strong agreement to strong disagreement with the deterrence hypothesis. The value in between corresponds to studies from which no usable opinion could be extracted or the degree of agreement depends heavily on different conditions. If an author does not distinguish between these crime categories, we assign his opinion to all three crime categories. From these variables, we generated an overall rating index O_{all} in the following way:

$$O_{\text{all},j} := \begin{cases} \left\lfloor \sum O_{cj} / \sum 1_{O_{cj}} \right\rfloor, & \text{if } \left(\sum O_{cj}\right) / \left(\sum 1_{O_{cj}}\right) < 0; \\ \left\lceil \sum O_{cj} / \sum 1_{O_{cj}} \right\rceil, & \text{if } \left(\sum O_{cj}\right) / \left(\sum 1_{O_{cj}}\right) > 0; \\ 0, & \text{if } \left(\sum O_{cj}\right) / \left(\sum 1_{O_{cj}}\right) = 0. \end{cases}$$

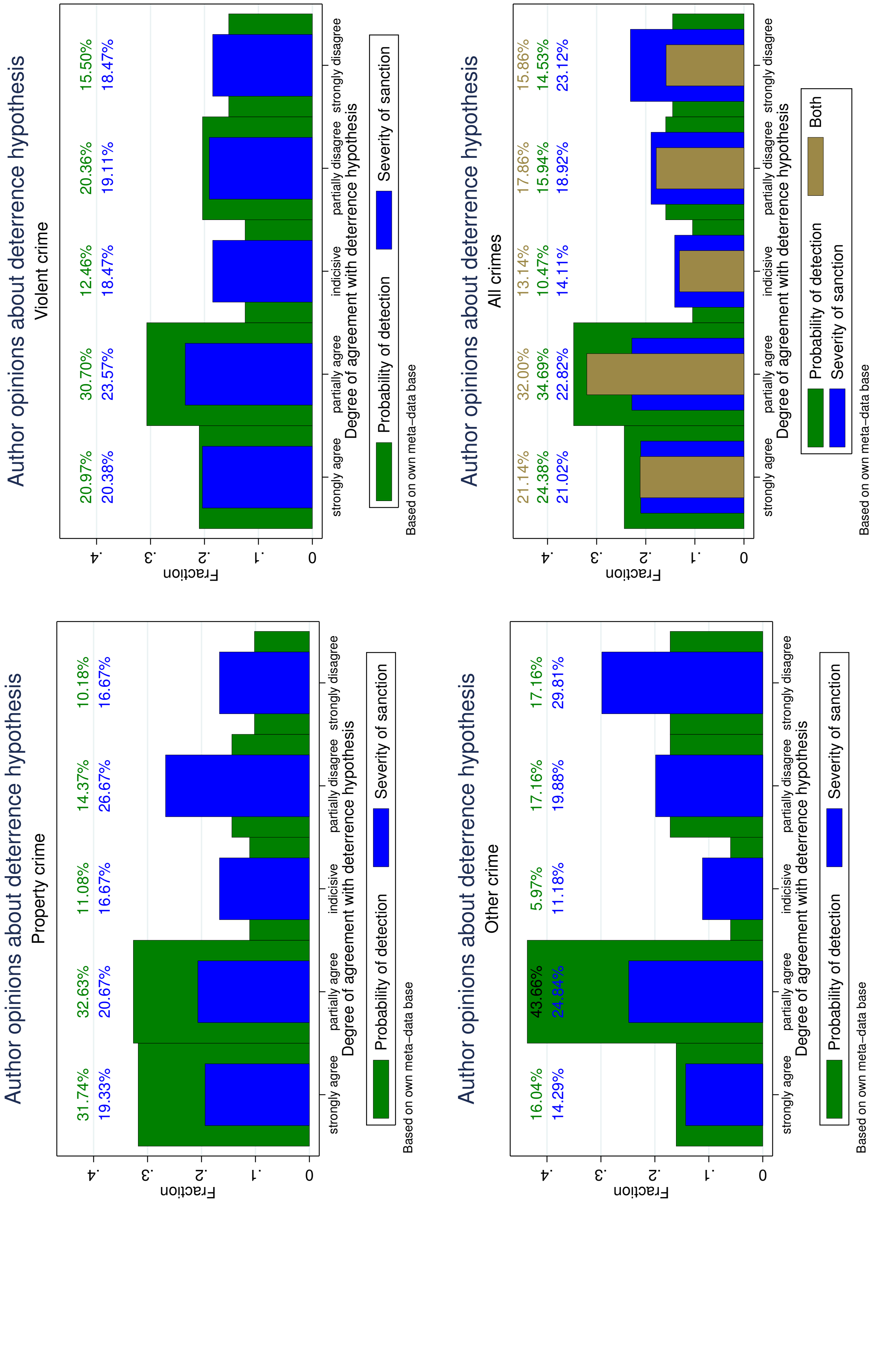
Whereas the sums are over the associated crime categories c for each study j. This index, and that for each crime category, can be seen in figure 3.5.

The first obvious observation is that the agreement with the deterrence hypothesis is consistently weaker in the case of the severity of punishment. This holds for all three crime categories; whilst being the weakest in the case of other crimes, followed by property crimes. Surprisingly, the best relationship between crime and the severity of punishment is found for violent crimes, although, these crimes are least affected by rational considerations (at least in theory). This may result from the more severe punishment (longer sentences, prison instead of fines, etc.) for violent crimes compared with property or other offenses. Regarding the probability of punishment, the picture is more intuitive: the agreement is strong for property crimes, while it is ambiguous for the other two categories. The partial agreement in the case of violent and other crimes is much stronger than the disagreeing opinions, while the strong agreement is on the same level as the disagreement (in regard to the probability). Additionally, the partial agreement is always the strongest; only in the case of property crime the strong agreement is almost equally large. In the case of the severity of punishment the opinions are almost evenly spread with the exception of "other" offenses: although the columns show a similar behavior, the disagreement is somewhat stronger.

Aggregating over all three crime categories, the authors do, more or less, agree with the deterrence theory in the case of the probability of punishment. When we aggregate over crime categories and kinds of punishment, the agreement outweighs disagreement by a large margin. This is based on the fact that the probability of punishment is studied more often than the severity of punishment[33].

[32] We are aware of the fact that the interpretation of the authors' statements by the reader may be problematic.
[33] In most cases this simply follows from the unavailability of appropriate data.

Figure 3.5: Opinion of the authors about the deterrence hypothesis



3.4 Publication Bias

> *The econometric art as it is practiced at the computer terminal involves fitting many, perhaps thousands, of statistical models. One or several that the researcher finds pleasing are selected for reporting purposes. This searching for a model is often well intentioned, but there can be no doubt that such a specification search invalidates the traditional theories of inference. [. . .] all the concepts of traditional theory, utterly lose their meaning by the time an applied researcher pulls from the bramble of computer output the one thorn of a model he likes best, the one he chooses to portray as a rose.*
>
> *Leamer (1983)*

When analyzing data, writing and publishing a study, the reported results may be biased for various reasons:

- The researcher initiative bias. This bias is introduced if the researcher intentionally or negligently forces his results in one direction. This may happen by inappropriately cleaning the data, using misspecified models or choosing inappropriate methods for evaluation. Then the "true" distribution of estimates the data yields does systematically differ from the distribution of the published estimates (Glaeser, 2006).
- The publication bias. This bias is introduced by publishing only those results, which seem to support some hypothesis and holding back those results, which do not (or vice versa). The distribution of the published estimates differs systematically from the distribution of the calculated estimates.
- Any other unintentional or unavoidable bias. Even if the author has done everything possible his estimates may still differ systematically from the distribution of the "true" estimates. This may be caused by the data itself, lack of knowledge about better methods, missing important variables and other reasons.

It is not surprising that different results can be drawn from the same data source, when enough free parameters are available, as Dijkstra (1995) puts it: "by simply adding one regressor one can obtain essentially every set of desired regression coefficients and predictions as well as t-values and standard errors."

Detecting a bias of the third category is merely of theoretical relevance and infeasible to measure in practice. Therefore, we will deal with the first and second category only. These are are undistinguishable to us since we have only access to the published studies. Subsequently we subsume these three categories in the following as publication bias. As McManus (1985) remarks, it is only natural and understandably that a researcher picks these specifications which he thinks are the best, “those that make the strongest case for the researcher’s prior hypothesis”. Hopefully, publication bias arises out of the good intention to report *the results he thinks best*, and not to mislead the public by reporting *the results he likes best*. Donohue and Wolfers (2005) emphasize this by remarking that such a bias “may occur without any of the authors being aware of it: they might simply want to report useful findings, and evidence falsifying a null hypothesis is typically regarded as more valuable”.

Publication bias as a problem has long been recognized by medical researchers and social scientists (Sterling, 1959; Rosenthal, 1979) and has also become popular in empirical economics (de Long and Lang, 1992). Traditionally, it is assumed that researchers, reviewers, editors and even readers are more willingly to accept positive results - i.e., the significant rejection of the null hypothesis of no effect (Stanley, 2005a). The most important reason to expect this, are that reviewers and editors might more readily accept results which are consistent with conventional views; only models are selected which show expected characteristics; researchers might want to get results consistent with their theory; everyone is more confident with significant results than inconclusive statistics; inadequate techniques might lead to (in)significant results or that suspicious data (outliers) may be wrongly excluded from a study.

Furthermore, the published results may tend to be more significant than they ought to be, since results - or even whole studies - which are insignificant may remain in the ’file-drawer’[34]. It has to be remembered that such omitted results are assumed to be insignificant - and not to have (as assumed by Rosenthal) an effect size or significance of zero (Scargle, 2000).

Usually, the publication process only refers to refereed journals and, as a consequence, attempts are made to minimize the publication bias by either requiring prior registration of studies (as is done by leading medical journals (Krakovsky, 2004)) or by including working papers and drafts in a meta analysis (e.g., Florax et al. (2002); Nijkamp and Poot (2005)). We think that neither method is sufficient for typical economic problems[35], since the main bias may be introduced during the conception and calculation of the estimations:

> *[. . .] there is much uncertainty as to the “correct” empirical model that should be used to draw inferences, and each researcher typically tries dozens, perhaps hundreds, of specifications before selecting one or a few to report.*
>
> *McManus (1985)*

Advances in technology have decreased the costs of running tests and alternative specifications. The consequence is that “the ability of researcher to influence results must be increasing over time” (Glaeser, 2006) and newer results should be faced with more scepticism. Techniques concerning meta analysis and publication bias which are used in the medical field are not applicable here,

[34] In fact, publication bias was initially called the ’file-drawer’ problem by Rosenthal (1978, 1979) and modified by Rosenberg (2005) and still refers to the intentional omitting of results (Scargle, 2000)

[35] This is especially the case in the field of empirical economics, where the researcher is free to chose his models and the estimation techniques. Furthermore, the “published” working paper version is usually already very close to the final version.

because they are too specialized (medical tests are almost always conceptualized as controlled experiments).

In theory, omitting insignificant results should be less of a problem in criminometrics, as pointed out by Eide et al. (1994), since evidence supporting or rejecting the deterrence hypothesis should be of equivalent importance - as is the case in some economic theories like the natural rate hypothesis (Stanley, 2005b). Nonetheless, authors preferring the deterrence hypothesis might omit insignificant or positively signed results, while those authors who do not like the deterrence hypothesis, might do the contrary. Insignificant results may be disliked by both types of authors: although insignificance already implies the absence of an affect, many people who oppose the deterrence hypothesis seem to "prefer" positive and significant results to discard any deterrent effect. Therefore, no reasonable prior assumption about the properties of the potential omitted studies can be made and methods based on the file-drawer approach should not be applied here. Stanley (2005a) advises against this approach anyway.

There is also a self-cleaning effect in every field of research: competition. A published study may be an incentive for other researchers to refute its results - even if it would have not been of any interest on its own (Glaeser, 2006). In the deterrence literature, there certainly was and is a strong scientific competition going on. As a consequence, increased scepticism of sceptic results is appropriate.

3.4.1 Methods to Detect Publication Bias

People must not attempt to impose their own truth on others.
The right to profess the truth must always be upheld, but not in a way that involves contempt for those who may think differently.
Truth imposes itself solely by the force of its own truth.

Karol Wojtyla (Pope John Paul II), 1991

Some authors try to circumvent publication bias by interpreting "missing" studies as "missing values" and augment them (Smith et al., 1997), but this seems unfeasible in our case. Most standard graphical methods, as summarized in Stanley (2005a), are also not applicable here, since they are based on the interpretation of graphs which are not compatible to our weighting scheme. However, we present some other graphs in subsection 3.4.2.

The principle of the main method we employ in subsection 3.4.3 is rather simple. In the case that an effect exists, the significance value should increase as the sample size increases (and the standard deviation decreases), whereas it should be independent of the sample size in the case of no effect. Leamer (1983) already pointed out that any null hypothesis can be rejected - whether reasonable or not - if the sample size is large enough. Usually, we don't expect this relationship to hold perfectly but it should, at least, be positive and significant when an effect exists (Stanley, 2005a). This test is used in many studies (see, for example, Stanley (2005a) and Waldorf and Byun (2005)).

Case I - Effect Does Not Exist

When there is no underlying effect, which implies that the null hypothesis $\beta = 0$ is true, then $t =\sim t(0, \nu)$ $\quad (= \beta/s$ in the standard regression analysis) and

$$\mathbb{E}[\log|t|] = \frac{\log\nu - \Psi(\frac{\nu}{2}) - \gamma - 2\log 2}{2},$$

whereas ν is the number of degrees of freedom, Ψ is the DiGamma-function and γ is the Euler-constant.

This immediately leads, as explained in section A.2, to $\mathbb{E}[\log|t|] = -(2\log 2 + \gamma)/2 + O(1/\nu)$. Thus, regressing $\log(|t|)$ on α_0 and $\alpha_1 \log\nu$ should result in $\alpha_0 \approx -0.635$ and $\alpha_1 = 0$.

Case II - Effect Does Exist

It it obvious that, if there is an effect, the t-value must increase as the sample size increases, because the estimation error decreases. Actually, if the hypothesis $\beta = 0$ is false, the t-value will follow the non-central t-distribution. As described in Stanley (2005a), footnote 10, $\mathbb{E}[|t|] = \frac{|\beta|}{\sigma_b}$ where $\sigma_b := \sigma[(X^TX)_b^{-1}]^{\frac{1}{2}}$ is the b-th diagonal element. Since the limit of X^TX/n is a finite, positive definite matrix σ_b is proportional[36] to $1/\sqrt{n}$.

Thus, regressing $\log(|t|)$ on α_0 and $\alpha_1 \log n$ should result in $\alpha_1 = 1/2$ and some α_0 (its value is not important).

3.4.2 Visual Analysis

Figuring things out for yourself is the only freedom anyone really has. Use that freedom.

Jean Rasczak in Starship Troopers*, 1997*

As mentioned before, our specific weighting scheme and the lack of a sufficient number of standard deviations, render many standard methods unusable (for example, a Galbraith plot is not compatible with weighting). Nevertheless, some plots are still helpful to judge how influential a publication bias may be.

A simple way to detect an obvious bias is to plot the histogram of the absolute (normalized) t-values. Prominent breaks of the distribution at the usual levels of significance are indicators of a publication bias. Another way to detect some misbehavior of the distribution of significance values is to standardize all (normalized) t-values within each study, and then to plot these standardized values. This procedure should make the variables more comparable by taking out the different levels and variations of each study. Any major breaks in the distribution of these standardized values can be interpreted as a indicator of publication bias.

[36] Since the diagonal elements of X^TX increase in n, those of $(X^TX)^{-1}$ decrease in n.

Since a histogram plot is not compatible with our analytic weights (w^a), our weighting scheme is approximated by the following frequency weights:

$$w_i^f = \text{round}\left(\frac{w_i^a}{\min(w^a)}\right).$$

Besides expanding the numbers of observations to meaningless values, deviations in the basic statistics are only introduced by rounding (the median remains the same, the first three moments change only by 0-2.5%), as shown in table 3.20.

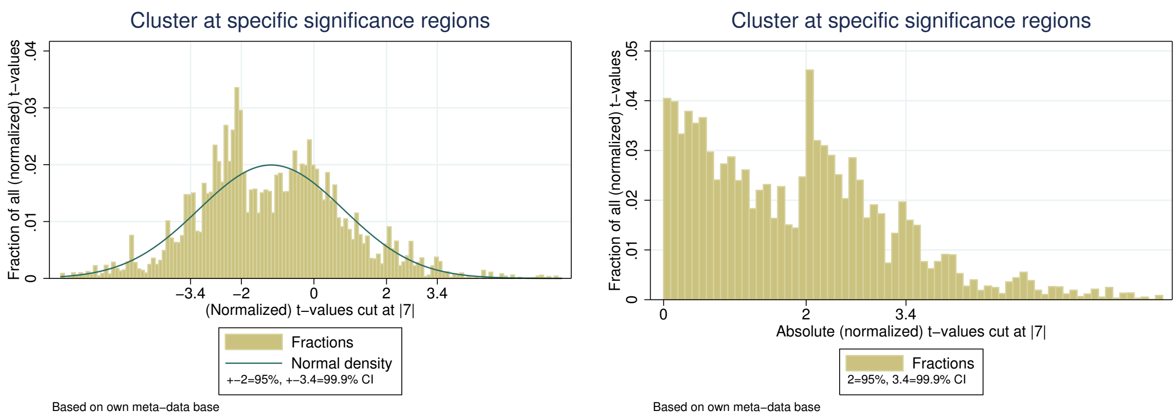
Table 3.20: Comparison of the (normalized) t-values with analytic and frequency weights

Weighting	mean	median	variance	skewness	kurtosis	n
Analytic	-1.404	-1.368	9.453	-1.948	31.210	6530
Frequency	-1.402	-1.368	9.463	-1.982	31.862	135227

n is the number of observations; in the case of the frequency weights it is inflated by multiplying the analytic weights by the inverse of their smallest value to achieve whole numbers.

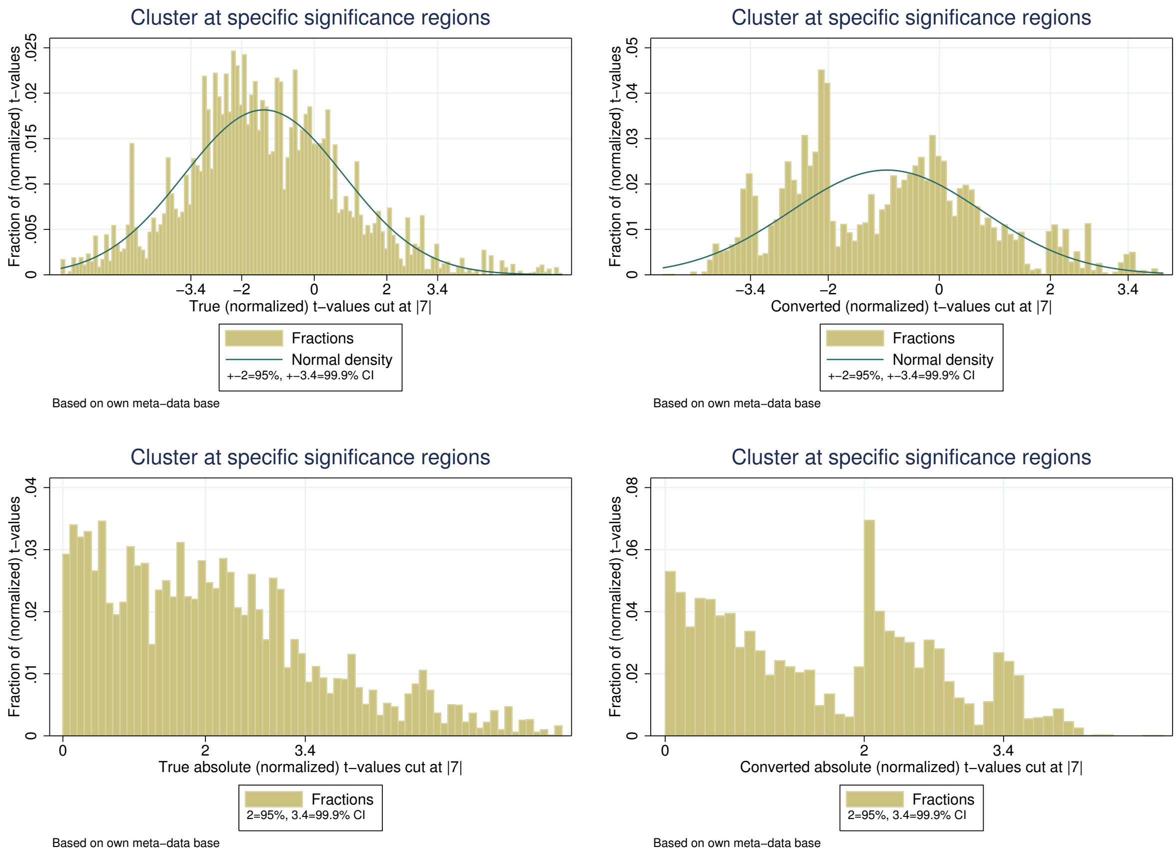
Usually the sign of the significance value is neglected and just the absolute values are considered. This procedure is correct and enhances the visual interpretation as long as the bias occurs for both signs in a similar fashion. Since it may be argued that positive (or negative) estimates of deterrent effects suffer more from a bias, we always consider both cases. The corresponding histograms can be found in figure 3.6.

Figure 3.6: Histograms of all (normalized) t-values



Although the spikes at the usual levels of significance are obvious, they cannot be interpreted as easily due to the transformation procedure from p-values into t-values (see subsection 3.2.2). Since the distribution of the true t-values (i.e., which were given by the study) does not show such a conspicuous behavior, as can be seen in figure 3.7, these spikes are largely based on the transformed p-values. How much this process distributes to these differences, cannot be judged unambiguously from these pictures - the number of significant p-values seems much to high (or the number of insignificant p-values is too low). Or to put it differently, the portion of insignificant p-values does not match those given by the (true) t-values.

Figure 3.7: Difference between the true and transformed (normalized) t-values



Looking at the basic statistics in table 3.21 of the true and transformed (normalized) t-values, we see that they are indeed quite differently distributed.

Table 3.21: Comparing statistics of the true and transformed (normalized) t-values

Variable	mean	median	sd	min	max	n	%w
True t	−1.7865	−1.6744	3.8155	−64.8110	19.0480	2811	54.35
Transformed t	−0.9481	−0.8892	1.7308	−4.9732	3.9604	3719	45.65

n is the number of observations, *%w* is the fraction of the summed weights of the corresponding category. “true” t-values are those which (or all necessary ingredients) are given explicitly by a study while “transformed” t-values are calculated or approximated by p-values or other sources.

Regarding the histograms, we have not taken into account the clustered structure (the values are not independent within studies and their levels may be very different across studies) of our data.

One way to mitigate the problem is to locally normalize the t-values:

$$t^l_{j,i} := \frac{t_{j,i} - \overline{t_i}}{\sigma_{t_i}},$$

whereas $t_{j,i}$ is the (normalized) t-value j from study i and t_i are all (normalized) t-values from study i. Aside from studies with only a single estimate, we have also to remove studies with two estimates, because $t^l_{1,i}$ (and $t^l_{2,i}$) evaluates to $\text{sign}(t_{1,i} - t_{2,i})\sqrt{0.5}$, independent of the actual (normalized) t-values.

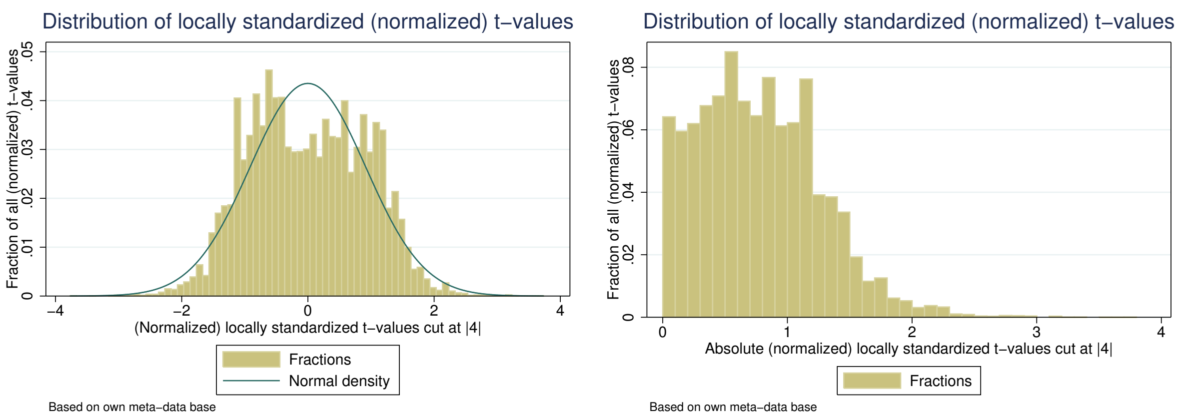
Dropping the studies with one or two estimates, we lose 284 observations (193 studies representing 27.76% of all weighted data). We then calculate the normalized values within each study, which should better represent the clustered structure. Table 3.22 gives the summary statistics of the local (normalized) values and figure 3.8 shows the histograms of these locally standardized (normalized) t-values.

Table 3.22: Statistics of the locally standardized (normalized) t-values

Variable	mean	median	sd	min	max	n
Mean t	−1.2652	−1.1870	2.0107	−17.1570	9.5653	6246
Sd	1.7157	1.3129	1.5384	0.0279	13.3750	6246
Standardized t	0.0000	−0.0189	0.9180	−5.5969	4.3330	6246

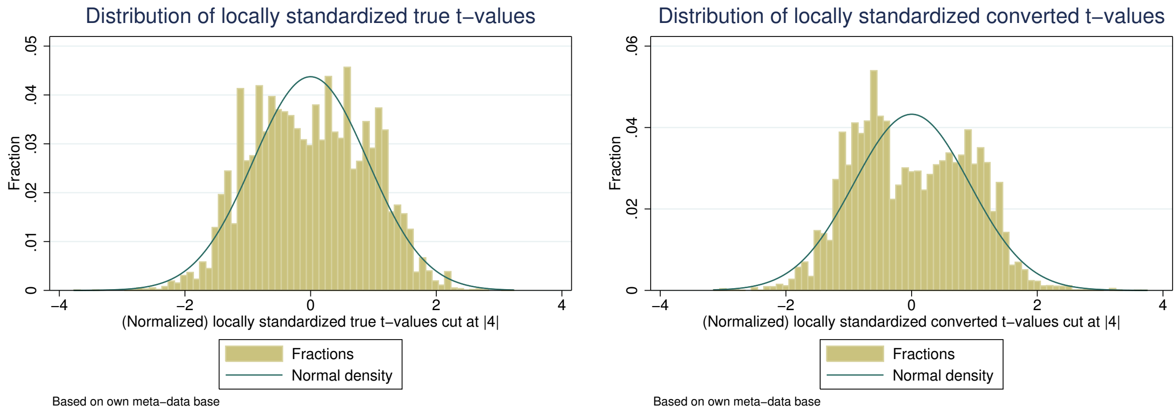
All values refer to the within-statistics of each study. Studies with only one or two estimates are dropped. n is the number of observations of the remaining 503 studies.

Figure 3.8: Histograms of all locally standardized (normalized) t-values



The sparse mid (and the flat beginning of the histogram) is obvious and it cannot be explained by the transformation because both kinds of t-values share, in principle, the same behavior (refer to figure 3.9). It is important to remember that values around zero represent the mean (normalized) t-values within each study.

Figure 3.9: Histograms of the true and transformed locally standardized (normalized) t-values



The visual analysis suggests that there are not enough values around the mean value of each study. This would explain both figures; that for all values, as well as the figure for the true and transformed values. Therefore, the conclusion of the visual analysis is that either many values of low or medium significance are missing or have partially been replaced by more significant - or very insignificant - values.

3.4.3 Analytical Analysis

I suppose it is tempting, if the only tool you have is a hammer,
to treat everything as if it were a nail.

Abraham H. Maslow, The Psychology of Science, 1966

In principal, we analyze the dependence of the absolute (normalized) t-values on the underlying sample size. The larger the sample size, the larger the value of significance should be if an effect exists. The following analysis is based on 5050 observations (representing 81.44% of the data) for which the sample size is available.

Regressing the logarithm of the absolute (normalized) t-values on the logarithm of the sample sizes (table 3.23), there is a significant and positive relationship to be found, although it is small. This can also be verified with the corresponding plot in figure 3.10. In all regressions we use the studies as clusters[37] because the observations within each study are certainly not independent. On the other hand, the ratio between clusters and observations is quiet small, so that we also show the results of a standard OLS regression.

Figure 3.10 indicates that the relationship largely depends on those estimates based on large sample sizes. Besides the sample size, there may be many other reasons driving the size and significance of the coefficient. Some of these are tested and reported in table 3.24.

[37]This means that the diagonal of the covariance matrix is only constant within each cluster.

Table 3.23: Regressing $\log(|t|)$ on $\log(n)$

Variable	coef.	robust sd	t	p
$\log(n)$	0.0448951	0.0203802	2.20	0.028***
Constant	0.0896568	0.1103612	0.81	0.417*

The regression is based on 5049 observation from 548 studies using clustered and robust standard errors (using the studies as clusters). $R^2 = 0.0045$, $F(1,547) = 4.85$, $P(F) = 0.028$, root mean square error (RMSE) is 1.1899. The symbols *,** and *** represent the significance in a regression without clustered standard errors at a 10, 5 and 1% level.

Figure 3.10: Relationship between $\log(t)$ and $\log(n)$

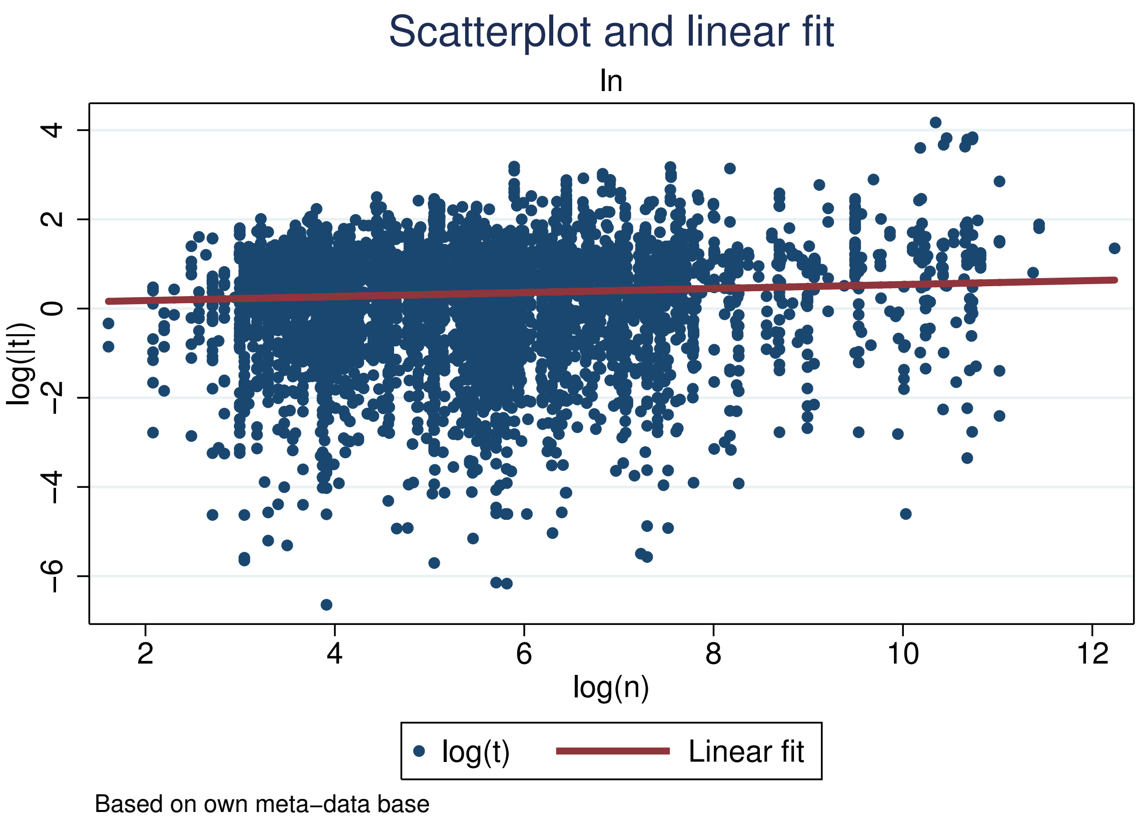


Table 3.24: Coefficients from regressing $\log(|t|)$ on $\log(n)$ for specific subsets

Subset	coef.	robust sd	t	p	n	%w	cluster
All	0.04490	0.02038	2.20	0.028***	5049	100.00	548
True t	0.02421	0.02503	0.97	0.334*	2371	57.77	329
Transformed t	0.09522	0.03196	2.98	0.003***	2678	42.23	248
$t < 0$	0.04328	0.02141	2.02	0.044***	3645	74.67	503

continued on the next page...

Subset	coef.	robust sd	t	p	n	%w	cluster
$t > 0$	0.06598	0.04147	1.59	0.113***	1404	25.33	299
$n \leq 60$	0.09354	0.14737	0.63	0.527	1305	26.52	159
$60 \leq n < 250$	−0.17363	0.14013	−1.24	0.217**	1219	26.40	162
$250 \leq n < 500$	0.85946	0.42491	2.02	0.046***	1025	16.21	108
$n \geq 500$	0.09770	0.05323	1.84	0.068***	1500	30.87	189
#est. ≤ 2	0.07476	0.04565	1.64	0.105	153	19.25	101
$3 \leq$ #est. ≤ 5	0.03804	0.05973	0.64	0.526	311	17.34	93
$6 \leq$ #est. ≤ 11	0.03503	0.04053	0.86	0.389	675	20.85	115
$12 \leq$ #est. ≤ 25	−0.00909	0.03643	−0.25	0.803	1248	23.53	129
#est. ≥ 26	0.09179	0.04224	2.17	0.032***	2662	19.03	110
S: Europe	−0.01965	0.07106	−0.28	0.783	736	12.87	69
S: USA	0.04787	0.02234	2.14	0.033***	3989	76.61	420
S: Other	0.11029	0.09963	1.11	0.273**	324	10.52	59
S: Economists	0.01007	0.02669	0.38	0.706	2147	48.11	265
S: Soc. & crim.	0.06384	0.03994	1.60	0.112***	2497	35.92	197
S: Others	0.17227	0.05384	3.20	0.002***	380	15.04	80
S: Journal	0.05602	0.02208	2.54	0.011***	4043	85.30	466
S: Book	−0.03650	0.07795	−0.47	0.643	362	5.48	30
S: Paper	−0.03076	0.05668	−0.54	0.590	644	9.22	52
S: Cross sections	0.00301	0.06189	0.05	0.961	1198	23.58	131
S: Time series	0.03773	0.03687	1.02	0.308**	1444	32.00	171
S: Panel data	0.09055	0.03958	2.29	0.024***	1006	17.26	96
S: Crime data	0.03956	0.02516	1.57	0.117***	3234	65.52	358
S: Surveys	0.14302	0.04778	2.99	0.003***	1735	27.58	155
S: Experiments	0.05778	0.08665	0.67	0.508	188	10.11	55
Crime data	0.04079	0.02781	1.47	0.143***	3048	61.32	340
Surveys	0.17936	0.05945	3.02	0.003***	1732	25.98	148
Experiments	0.06537	0.04650	1.41	0.164**	269	12.70	68
Bivariate	0.14258	0.03593	3.97	0.000***	1491	22.07	153
Multivariate	0.01938	0.02249	0.86	0.389*	3558	77.93	467

The first regression is based on 5049 observation from 548 studies using clustered and robust standard errors (using the studies as clusters). Some categories do not some up to one because of missing or multiple entries. Variables with a preceding "*S:*" are measured on the study-level while all others are based on the estimates. n is he number of observations in the corresponding category. $\%w$ is the fraction of the summed weights of the corresponding category. The symbols *,** and *** represent the significance in a regression without clustered standard errors at a 10 ,5 and 1% level.

end of the table 3.24

We have to be cautious when interpreting the results from table 3.24, because most coefficients are close to zero and the categories are based on different numbers of observations. Nevertheless, it seems to be that the data offers some clues. There is no evidence, if a publication bias is present, that positive or negative results are more affected. However, there are suspicious results when we look at studies with moderate sample sizes. It seems to be, consulting figure 3.10, that there are too many low absolute (normalized) t-values, with sample sizes between 250 and 500, than there should be. On the other hand, there seem to be too many large absolute values for results based on sample sizes between 60 and 250. The high coefficient for surveys may be based on the fact that the results scale more heavily with the sample size than other research methods do. For example, panel data is expected to yield much lower coefficients than Pearson correlations. On the other hand, panel data are usually based on many observations and, therefore, the significance of their results scale much better than in the case of cross sections or time series. The fact that journals have the "best" coefficient contradicts with the common hypothesis that publication bias is more readily found in (refereed) journals. The contrary is the case: results published in working papers and books, which are usually not subject to a referee process, have even the wrong sign but are not significant. So, if at all, these are biased towards zero. The relationship between the sample size and significance is more profound for bivariate methods. To some extent, this may be attributed to the dominant usage of Pearson correlations, multivariate methods which do not have t-distributed standard errors under the null hypothesis, or their complexity. The difference between economists on the one side and sociologists, criminologists and jurists may be partly explained by the fact that economists do use bivariate methods and surveys less often. There seems to be no relationship in the case of studies about Europe (mainly the United Kingdom and Germany). Whether this is based on some artifacts of the data, publication bias or a weaker deterrent effect cannot be unambiguously judged here. It is somewhat disappointing that no clear assertion can be made for the true t-values. This may be partly explained by the fact that t-values are much more often reported by economists who use multivariate methods and employ crime data.

As pointed out by Stanley (2005a), the absence of any relationship can lead to two different conclusions: either there is a publication bias present or there is simply no effect. The latter option is often left aside in the literature, leading to false conclusions and thus has to be considered separately. In the case of the deterrence literature this seems not to be the case, because there is an overall relationship and insignificant results for specific subsets should be based on a bias rather than on the absence of an effect.

Even if there is an effect, the regression itself cannot tell us the characteristics of the bias. Usually, there will be clusters around the typical regions of significance, which reduce the influence of the sample size. As shown in subsection 3.4.2, it seems to be that there is a shift away from "medium" insignificance in the distribution but no obvious clustering.

Other Evidence

We have also some other statistics at our disposal. We know whether the deterrence variable was the main focus of an estimation or just a covariate (e.g., in studies analyzing the effect of crime on unemployment). In principle, the mean effect should be independent of the focus of the estimation. However, the mean (normalized) t-value for studies which focus on deterrence is -1.47 (the median is -1.44), while the values for estimates which do not focus on deterrence, differ significantly (mean -1.02, median -1.01). This tentative evidence should not be overrated, because estimates, which do not focus on deterrence, may lack important features to minimize potential errors, biases and other specialities found in the deterrence literature.

3.4.4 Subsequent Consequences

Absence of evidence is not evidence of absence.

Carl E. Sagan, The Dragons of Eden, 1977

We could find only some evidence of publication bias. Since the data is very heterogenous, it is not surprising that the analytical analysis has not produced any large coefficients but most of them are correct (positive and significant). The main problem seems to be the reported results based on medium sample sizes. Results based on 60 to 250 observations are significantly biased towards zero, while those, based on 250 to 500 observations, show the opposite behavior. Overall, there seems to be a lack of results lying near the mean of each study - and these means are distributed around -1.27. This value might not be appreciated very much by researchers, because those results are neither clearly insignificant nor are they in range of the usual levels of significance. Therefore, it is not surprising that values from this region are found less often than expected. However, the calculated coefficients do not allow us to adjust the (normalized) t-values, as in Stanley (2005a), because we cannot identify their distribution. Even an author with no ill intentions might respecify his model until it looks well behaved, which may lead to the described sparse areas. In the multivariate analysis (see section 3.6), which controls for many model properties, this might not be that much of a problem, while the interpretation is more difficult for bivariate and overall statistics, like the average or median (normalized) t-value.

The consequences for this meta analysis is to keep this in mind and be careful. We cannot correct this potential publication bias, since there is no way to distinguish unbiased from biased estimates. The least we can do is to split the number of observations into two variables; each taking the value zero in the case of negative (positive) normalized t-values, to eliminate the technical influence of the sample size of each estimate. Nevertheless, we should take heed of the advise from Glaeser (2006) that “if we are faced with the choice between no information and biased information, the latter option is preferred”.

3.5 Bivariate Statistics

Now that we have all this useful information, it would be nice to do something with it. (Actually, it can be emotionally fulfilling just to get the information. This is usually only true, however, if you have the social life of a kumquat.)

Unix Programmer’s Manual

The extent of the support of the deterrence hypothesis varies largely between different groups. In this section we show several key statistics for some selected groups. These relationships should not be overrated, since correlations with other unaccounted variables may be the main reason for many dependencies (see section 2.3.1).

A first impression of the dependencies is given by the simple correlations between the (normalized) t-values and the available variables in table 3.25. The largest correlation coefficients belong, as expected, to variables related to the used sample size of the estimates. In the case a deterrent effect exists, the t-values scale with the sample size as explained in section 3.4. It is obvious that variables, which relate to economic authors, journals, studies and economic theories come all with quite negative coefficients, while those relating to the criminological field are all positive[38]. These coefficients are not surprising because it is commonly assumed that economists more readily agree with the deterrence hypothesis, while criminologists are rather sceptical about it. Another observation is that all significant coefficients relating to experiments are negative as well. It is a curious result that reporting tests of significance is negatively correlated with the (normalized) t-values on the study-level but positively on the estimate-level. Estimates, which are based on simple correlations seem to yield more positive values, while logit (and probit) estimates are negatively correlated. It is also interesting that using the strength of the police yields more positive values, similar to those estimates which do not focus on deterrence. Newer studies, working papers and reports, as well as those estimates based on youth-samples come along with rather negative (normalized) t-values.

In the subsequent subsections, most tables show the mean and median (normalized) t-value, as well as the percentage of estimates which are consistent with the deterrence hypothesis and are significant at a 5% level. All entries are sorted by their median. For each group an Analysis of Variance (ANOVA) is performed with the most frequent entry as the reference category. All entries which are not significant at a 5% level, but are interesting for some other reasons, are written in *italic*.

3.5.1 Study-Related Groups

Although to penetrate into the intimate mysteries of nature and thence to learn the true causes of phenomena is not allowed to us, nevertheless it can happen that a certain fictive hypothesis may suffice for explaining many phenomena.

Leonhard Euler, Introductio in analysin infinitorum, 1748

In the following, we will present the dependency within certain study-related groups. This means that the grouping variable is constant within each study. We inspect the relationship between several interesting variables and the (normalized) t-values and comment on it.

As depicted in table 3.26, compared to U.S. authors, studies from authors of the most frequent and largest European countries (the United Kingdom and Germany) yield significantly better results (in favor of the deterrence hypothesis). Their medians are -1.85 and -1.83, which is almost

[38] The user tr recorded all economic studies, the user aw most of the criminological and sociological studies.

Table 3.25: Significant correlations with the (normalized) t-values

Variable	coef.	variable	coef.
Study: size of first realized sample	$-0.235^{\star\star}$	Study: publication, year	-0.056
Study: size of second realized sample	$-0.119^{\star\star}$	Study: author, Isaac Ehrlich	-0.056
Study: user, tr	$-0.113^{\star\star}$	Study: complete sample	-0.055
Study: publication, economics	$-0.110^{\star\star}$	Estimate: exogenous, survey, other	-0.055
Study: error and plausibility checks	$-0.103^{\star\star}$	Study: journal, Review of Economics and Statistics	-0.053
Estimate: sub-sample of youths	$-0.103^{\star\star}$	Estimate: deterrence is focus-variable	-0.053
Study: institute, economics	$-0.099^{\star\star}$	Estimate: exogenous, binary category	-0.050
Study: journal, Economic Inquiry	$-0.097^{\star\star}$	Study: cross section	$+0.050$
Estimate: number of observations	$-0.093^{\star\star}$	Study: sample individuals, first population, pupils	$+0.051$
Estimate: covariate, Fixed effects (spatial)	$-0.092^{\star\star}$	Study: institute, criminology	$+0.051$
Study: author, economics	$-0.092^{\star\star}$	Estimate: exogenous, not in logs	$+0.052$
Study: economic, rational choice theory	$-0.091^{\star\star}$	Estimate: exogenous, death penalty, existence of death penalty	$+0.052$
Study: main location > 500000 inhabitants	$-0.091^{\star\star}$	Study: journal, Criminology	$+0.053$
Study: rate of return of second sample	$-0.088^{\star\star}$	Estimate: deterrence is covariate	$+0.053$
Study: experiment (laboratory)	$-0.085^{\star\star}$	Study: author, criminology	$+0.055$
Estimate: exogenous, experiment, experimental variation of probability of detection	$-0.081^{\star\star}$	Study: user, aw	$+0.055$
Study: author, psychology	$-0.080^{\star\star}$	Estimate: exogenous, death penalty, execution rate	$+0.055$
Study: sample individuals, second population, miscellaneous	$-0.076^{\star\star}$	Study: author, William C. Bailey	$+0.056$
Estimate: exogenous, experiment, yes	$-0.076^{\star\star}$	Estimate: multivariate method, path analysis	$+0.057$
Estimate: covariate, marital status	$-0.076^{\star\star}$	Estimate: covariate, urbanity	$+0.061^{\star}$
Estimate: exogenous, experiment, relates to the present	$-0.075^{\star\star}$	Study: author, sociology	$+0.063^{\star}$
Estimate: exogenous, crime data, arrest rate	$-0.074^{\star\star}$	Estimate: bivariate method, correlation	$+0.065^{\star}$
Estimate: covariate, personal characteristics	$-0.073^{\star\star}$	Estimate: covariate, previous convictions	$+0.066^{\star}$
Study: institute, psychology	$-0.072^{\star}$	Estimate: exogenous, crime data, police expenditures	$+0.067^{\star}$
Estimate: exogenous, other transformation	$-0.068^{\star}$	Estimate: test of significance	$+0.067^{\star}$
Study: sample unit, second population, individuals	$-0.067^{\star}$	Study: institute, miscellaneous	$+0.069^{\star}$
Study: experimental	$-0.066^{\star}$	Study: sample unit, first population, states	$+0.071^{\star}$
Estimate: endogenous, other	$-0.066^{\star}$	Study: journal, Criminal Justice	$+0.072^{\star}$
Study: sample unit, first population, individuals	$-0.066^{\star}$	Estimate: covariate, poverty, welfare	$+0.073^{\star\star}$
Estimate: exogenous, in logs	$-0.065^{\star}$	Estimate: study type, death penalty	$+0.079^{\star\star}$
Study: publication, working paper, report	$-0.064^{\star}$	Study: not experimental	$+0.082^{\star\star}$
Study: tests of significance	$-0.061^{\star}$	Estimate: weighted model	$+0.088^{\star\star}$
Study: sample base, second population, complete country	$-0.060^{\star}$	Study: traditional theory	$+0.111^{\star\star}$
Estimate: exogenous, crime data, conviction rate	$-0.059^{\star}$	Estimate: exogenous, crime data, police strength	$+0.126^{\star\star}$
Estimate: multivariate method, logit, probit	-0.058	Study: publication, criminology	$+0.130^{\star\star}$
Study: first population, United Kingdom	$-0.057^{\star}$		

A correlation coefficient is listed if its absolute value is larger than 0.05, is significant at the 0.01% level (two-sided test) and varies in at least 1% of the data. * marks p-values which are smaller than $5 \cdot 10^{-6}$ and ** those below $5 \cdot 10^{-9}$.

40% smaller than those of U.S. authors. Contrarily, Canadian authors find deterrent effects in much fewer cases (the percentage[39] is reduced by one third, compared with those authors from the UK and Germany). Estimates from Australian authors do not significantly differ from those of U.S. authors. The statistics of the authors from less frequent countries ("other") are also interesting: the mean and median estimate is much more negative than those from U.S. authors, while the percentage is much lower. This could mean that the results from those authors are more concentrated in the negative "no man's land", which is rather uncommon in our meta-data base (refer to section 3.4 about publication bias).

Table 3.26: Differences by the authors' nationality

Nation	mean	median	%	#e	#s
Finland	−3.03	−2.92	80.00	50	5
Israel	−1.64	−2.15	53.60	85	9
UK	−1.87	−1.85	48.76	275	28
Germany	−1.86	−1.83	42.15	208	22
Netherlands	−1.11	−1.46	43.58	81	8
Other	−1.81	−1.44	34.91	239	24
Overall mean	−1.40	−1.37	41.66	6530	663
USA	−1.38	−1.29	41.90	5143	522
Australia	−1.30	−1.01	44.87	132	13
Switzerland	−0.80	−0.99	38.81	66	7
Canada	−0.86	−0.89	31.22	336	34
Sweden	0.15	−0.15	7.87	61	6

Mean and median correspond to the (normalized) t-values of the particular group. % is the percentage of estimates which are consistent with the deterrence hypothesis and significant at a 5% level in a two sided test. *#e* is the weighted number of all valid estimates. *#s* is the number of studies the estimates are based on. *Italic* entries are not significant in the ANOVA at a 5% level, but are included, like the overall mean, to show otherwise interesting or some selected groups. The underlined entry is the reference category.

As mentioned before, it is almost common knowledge that economists more readily agree with the deterrence hypothesis than criminologists and sociologists. This view is supported by table 3.27. Psychology is, undisputable, number one in that list with a median of −2.15 and a percentage of 65.59%. These authors mostly study alcohol related offenses which yield rather negative results. Economics is the second most dominant category in favor of deterrence; the estimates have a median of −1.67 and 43.82% are consistent with the deterrence hypothesis and significant. As expected, sociologists show significantly less results in agreement with deterrence (median −1.01, percentage 38.67%). With a median of just −0.62 and a percentage of 33.31%, criminologists are at the very bottom of that list. Authors from "other" disciplines, which are more or less not related to deterrence research (e.g., mathematics, medicine, etc.), produce estimates which are very similar to the overall mean.

[39]In this context, "percentage" always refers to the percentage of (normalized) t-values which are consistent with the

Table 3.27: Differences by the authors' discipline

Discipline	mean	median	%	#e	#s
Psychology	−2.58	−2.15	65.59	275	27
Economics	−1.73	−1.67	43.82	2807	287
Overall Mean	−1.40	−1.37	41.71	6101	617
Other	−1.17	−1.33	41.25	767	76
Law	−1.17	−1.11	38.53	233	23
Sociology	−1.07	−1.01	38.67	1690	170
Criminology	−0.93	−0.62	33.31	738	75

Mean and median correspond to the (normalized) t-values of the particular group. % is the percentage of estimates which are consistent with the deterrence hypothesis and significant at a 5% level in a two sided test. *#e* is the weighted number of all valid estimates. *#s* is the number of studies the estimates are based on. *Italic* entries are not significant in the ANOVA at a 5% level, but are included, like the overall mean, to show otherwise interesting or some selected groups. The underlined entry is the reference category.

As stated in section 3.4 about publication bias, results may be different for various types of publication. Journals make up for the most studies we included (table 3.28). Working papers (including discussion papers, official reports, etc.) is the only category in which the (normalized) t-values are significantly different (the mean of −2.26 is quite smaller than −1.34 from the journals); but this difference is almost nullified when we look at the median or the percentage (−1.37 and 43.43% compared to −1.35 and 41.74%). All in all, there seems to be no major differences between the various types of publication; only that the average (normalized) t-value is more negative (−1.93) for books (the median of 1.57 and the percentage of 46.38% are different, but not significantly).

Table 3.28: Differences by the type of publication

Type	mean	median	%	#e	#s
Edited volume	−1.62	−1.63	43.82	261	28
Book	−1.93	−1.57	46.38	98	10
Other	−1.19	−1.41	31.89	214	22
Working paper	−2.26	−1.37	43.43	322	34
Overall mean	−1.40	−1.37	41.66	6530	663
Journal	−1.34	−1.35	41.74	5635	569

Mean and median correspond to the (normalized) t-values of the particular group. % is the percentage of estimates which are consistent with the deterrence hypothesis and significant at a 5% level in a two sided test. *#e* is the weighted number of all valid estimates. *#s* is the number of studies the estimates are based on. *Italic* entries are not significant in the ANOVA at a 5% level, but are included, like the overall mean, to show otherwise interesting or some selected groups. The underlined entry is the reference category.

Most studies are published in five countries, dominated by U.S., as shown in table 3.29. How-

deterrence hypothesis and significant at a 5% level in a two-sided test.

ever, there are almost no differences. Only the Netherlands stick out with a very low percentage (28.99%) and an average (normalized) t-value near zero (−0.46). An interesting observation is that studies in Canadian publications report more negative (normalized) t-values (median −1.67) than the average, while authors from Canada report less negative (−0.89) values.

Table 3.29: Differences by the country of publication

Country	mean	median	%	#e	#s
UK	−1.53	−1.82	47.57	539	54
Canada	−1.42	−1.67	39.37	234	23
Overall mean	−1.40	−1.37	41.66	6530	663
Germany	−1.50	−1.34	36.92	299	31
<u>USA</u>	−1.41	−1.29	41.84	5101	518
Netherlands	−0.46	−0.96	28.99	166	17
Other	−1.59	−0.94	41.16	190	20

Mean and median correspond to the (normalized) t-values of the particular group. % is the percentage of estimates which are consistent with the deterrence hypothesis and significant at a 5% level in a two sided test. *#e* is the weighted number of all valid estimates. *#s* is the number of studies the estimates are based on. *Italic* entries are not significant in the ANOVA at a 5% level, but are included, like the overall mean, to show otherwise interesting or some selected groups. The <u>underlined</u> entry is the reference category.

We distinguish the discipline of the publisher in table 3.30. Economic (and psychologic) publications are much more supportive of the deterrence theory (−2.15 and −1.73, 67.45% and 45.50%), while sociological (−1.33, 40.64%) are significantly less supportive. Criminological publications appear, after a large gap, at the very bottom of the list (−0.51, 30.85%). It is not astounding that there are no major differences to table 3.27 because sociologists and criminologists only rarely publish in economic media (as illustrated in table 3.9).

Table 3.30: Differences by the publishers' discipline

Discipline	mean	median	%	#e	#s
Psychology	−2.12	−2.15	67.45	183	18
<u>Economics</u>	−1.86	−1.73	45.50	2348	241
Other	−1.54	−1.42	45.24	945	96
Overall mean	−1.40	−1.37	41.70	6426	651
Sociology	−1.15	−1.33	40.64	1267	128
Law	−1.56	−1.20	38.25	346	34
Criminology	−0.62	−0.51	30.85	1335	134

Mean and median correspond to the (normalized) t-values of the particular group. % is the percentage of estimates which are consistent with the deterrence hypothesis and significant at a 5% level in a two sided test. *#e* is the weighted number of all valid estimates. *#s* is the number of studies the estimates are based on. *Italic* entries are not significant in the ANOVA at a 5% level, but are included, like the overall mean, to show otherwise interesting or some selected groups. The <u>underlined</u> entry is the reference category.

In table 3.26, we have seen the relationship with the countries the authors worked in. Although many authors study data of their own country, there are some differences. Looking at table 3.31, which shows the statistics diversified by the studied countries, some countries remain fairly stable in their position: Finland remains at the top (median −2.65, percentage 83.40%), while the UK, Germany and the USA remain at their positions as well. Australia switches from below the mean almost to the top (from −1.01 to −2.24 and from 44.87% to 64.39%). Sweden remains at the lower part of the table, but has now a more reasonable (higher) percentage (7.87% to 24.65%). Studies using Canadian data have estimates which are in least agreement with the deterrence hypothesis (−0.52, 28.26%). Although some crimes are studied more often in a country than in another (e.g., drunken driving in scandinavian countries or marijuana consumption in Australia), there is no obvious explanation why results based on Canadian data are in "worst" compliance with the deterrence hypothesis.

Table 3.31: Differences by the studied nation

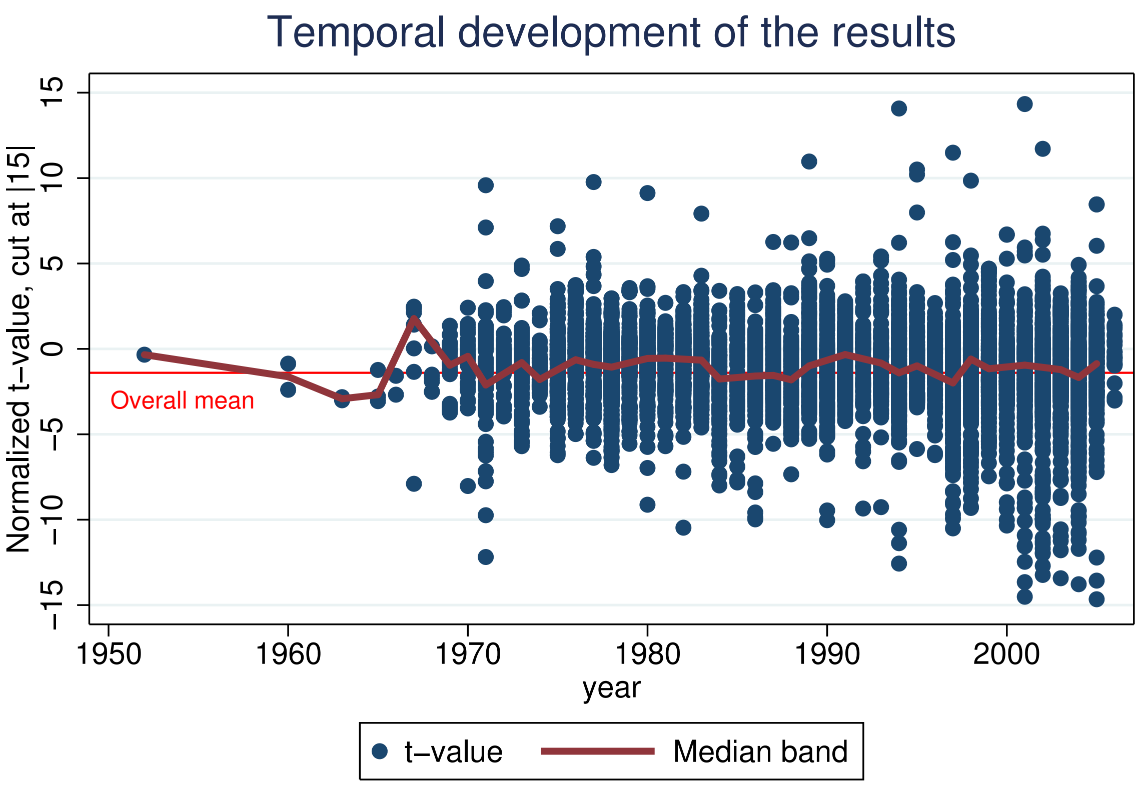
Nation	mean	median	%	#e	#s
Finland	−2.93	−2.65	83.40	60	6
Australia	−1.74	−2.24	64.39	112	11
UK	−2.17	−2.02	54.07	327	33
Germany	−1.96	−1.83	42.24	177	19
Netherlands	−1.79	−1.37	41.98	81	8
Overall mean	−1.40	−1.37	41.66	6530	663
<u>USA</u>	−1.37	−1.30	41.84	5008	507
Other	−1.37	−1.19	31.88	407	43
Switzerland	−0.97	−1.03	38.81	66	7
Sweden	−1.31	−0.69	24.65	92	9
Canada	−0.70	−0.52	28.26	280	28

Mean and median correspond to the (normalized) t-values of the particular group. % is the percentage of estimates which are consistent with the deterrence hypothesis and significant at a 5% level in a two sided test. *#e* is the weighted number of all valid estimates. *#s* is the number of studies the estimates are based on. *Italic* entries are not significant in the ANOVA at a 5% level, but are included, like the overall mean, to show otherwise interesting or some selected groups. The <u>underlined</u> entry is the reference category.

During the last 40 years much has or may have changed in the literature about deterrence: data, data quality, cultural backgrounds, estimation technology, attitudes of authors, the audience, offenders, and much more. Figure 3.11, however, reveals that there are no obvious time effects. Nevertheless, we have partitioned the estimates into five categories[40]. There are indeed some significant differences, as shown in table 3.32, but the differences are neither easily interpreted, nor are these very pronounced. If at all, there is a periodical pattern. However, these results are not robust when redefining the time categories.

[40] As always in such cases we chose the thresholds in such a way that each category contains an approximately equal number of observations.

Figure 3.11: Temporal development of the (normalized) t-values



Based on own meta data base

We make the same comparison with the utilized data. Table 3.33 shows that the median values differ as much as in the case of the year of publication; the percentages only partially. Although the entries are all significantly different from the newest data (median -0.92, 36.49%), there is a slight trend in the order of the rows. If we exclude the oldest data, studies using newer data seem to produce less significant results. Additionally, the ordering is somewhat different from that in table 3.32 because there is, if at all, a time trend instead of a periodic pattern.

The number of reported estimates certainly depends on the type of publication, because there is much more room to present results in books and working papers than in journals. The mean number of published estimates is 22 (median 8); the largest number is 764 (the smallest zero). The corresponding ANOVA can be found in table 3.34. Although not significant there is an obvious descending order with an increasing number of estimates. This may be, at least partially, explainable by technical reasons: presenting more results is commonly done for robustness checks (often by economists), which come along with more “contaminated[41]” estimates. Another reason is that simple correlation coefficients, which are associated with insignificant values (table 3.47), appear often in large numbers in a study.

[41]When evaluating numerous specifications a certain percentage may suffer from a misspecification bias.

Table 3.32: Differences by the year of publication

Years	mean	median	%	#e	#s
1979-1986	−1.42	−1.68	46.26	1225	125
1987-1994	−1.36	−1.53	43.05	1428	147
2001-2006	−1.90	−1.46	44.97	1412	141
Overall mean	−1.40	−1.37	41.66	6530	663
1952-1978	−1.12	−1.18	39.25	1237	125
1995-2000	−1.15	−1.06	34.06	1229	125

Mean and median correspond to the (normalized) t-values of the particular group. % is the percentage of estimates which are consistent with the deterrence hypothesis and significant at a 5% level in a two sided test. *#e* is the weighted number of all valid estimates. *#s* is the number of studies the estimates are based on. *Italic* entries are not significant in the ANOVA at a 5% level, but are included, like the overall mean, to show otherwise interesting or some selected groups. The underlined entry is the reference category.

Table 3.33: Differences by the year of utilized data

Years	mean	median	%	#e	#s
1966-1974	−1.11	−1.61	43.71	962	97
1975-1982	−1.17	−1.31	42.93	1023	103
1983-1988	−2.21	−1.20	41.73	824	85
Overall mean	−1.33	−1.19	39.25	4862	494
1875-1965	−1.05	−0.94	32.07	998	103
1989-2004	−1.25	−0.92	36.49	1056	106

Mean and median correspond to the (normalized) t-values of the particular group. % is the percentage of estimates which are consistent with the deterrence hypothesis and significant at a 5% level in a two sided test. *#e* is the weighted number of all valid estimates. *#s* is the number of studies the estimates are based on. *Italic* entries are not significant in the ANOVA at a 5% level, but are included, like the overall mean, to show otherwise interesting or some selected groups. The underlined entry is the reference category.

It may be reasoned that the design of a study may lead to different results. It is obvious that experiments come along with very negative (normalized) t-values (−2.1 to −2.45, 58.57% to 65%), as table 3.35 shows. The effect of the other designs are not that obvious. Surveys appear at the upper, middle and lower part of the table, depending on their design. Estimates based on reported crimes are below the mean. However, most of the non-experiments are not significantly different from studies based on time series (the reference category, −1.34, 41.69%).

Section 3.3 shows that 27.6% of all studies come from the top 22 authors (2.6% of all involved authors). It is reasonable to assume that individual preferences of the authors may lead to different estimates, depending on their personal attitude and other reasons. We stress that table 3.36 does not show any (clear) evidence for a publication or author bias; authors may prefer different methods, offenses, countries and other things which may lead to more or less significant estimates. The most striking result is that, among these 22 authors, the percentage of theory-consistent and

Table 3.34: Differences by the number of reported estimates

Number	mean	median	%	#e	#s
0-2	−1.58	−2.09	53.96	1349	133
Overall mean	−1.46	−1.41	42.82	5320	540
3-5	−1.25	−1.41	42.11	791	79
6-11	−1.29	−1.33	39.69	967	98
12-25	−1.60	−1.10	40.17	1021	104
26-764	−1.49	−1.00	35.49	1192	126

Mean and median correspond to the (normalized) t-values of the particular group. % is the percentage of estimates which are consistent with the deterrence hypothesis and significant at a 5% level in a two sided test. *#e* is the weighted number of all valid estimates. *#s* is the number of studies the estimates are based on. *Italic* entries are not significant in the ANOVA at a 5% level, but are included, like the overall mean, to show otherwise interesting or some selected groups. The underlined entry is the reference category.

Table 3.35: Differences by the design of a study

Design	mean	median	%	#e	#s
Experiment (by institution)	−2.55	−2.45	63.33	102	10
Experiment (laboratory)	−2.63	−2.36	58.57	285	28
Experiment (by researcher)	−1.03	−2.10	65.00	102	10
Survey (once)	−1.55	−1.60	45.63	1161	119
Experiment (natural)	−1.26	−1.60	44.90	197	20
Overall mean	−1.41	−1.37	41.69	6479	658
Time series	−1.39	−1.34	39.06	2235	224
Survey (panel)	−1.50	−1.32	44.52	367	38
Panel data	−1.57	−1.13	38.68	1000	101
Cross section	−1.12	−1.04	37.00	1517	156
Survey (multiple)	−1.46	−0.86	36.15	205	22

Mean and median correspond to the (normalized) t-values of the particular group. % is the percentage of estimates which are consistent with the deterrence hypothesis and significant at a 5% level in a two sided test. *#e* is the weighted number of all valid estimates. *#s* is the number of studies the estimates are based on. *Italic* entries are not significant in the ANOVA at a 5% level, but are included, like the overall mean, to show otherwise interesting or some selected groups. The underlined entry is the reference category.

significant results is almost doubled in the top five rows compared to the bottom five. While the upper position of Isaac Ehrlich (−3.34, 73.65%) and William C. Bailey (−0.26, 13.92%) in the last row is not surprising, it is rather curious to find Steven D. Levitt (−0.83, 32%) among the bottom five. As stated before, there are only two authors who did not live in the USA at the time of writing and both appear above the mean (−1.37, 41.66%); Matti Virén has the highest percentage (80%), while Horst Entorf[42] is slightly above the mean (42.96%).

[42] Although his entry is just below the mean and the reference group, the corresponding mean and percentage indicates that he has to be associated, with the "upper" part.

Table 3.36: Differences by prominent authors

Author	mean	median	%	#e	#s
Simon Hakim	−2.91	−3.47	63.82	67	7
Isaac Ehrlich	−3.13	−3.34	73.65	64	7
Matti Virén	−3.03	−2.92	80.00	50	5
Harold G. Grasmick	−2.14	−2.58	72.61	101	10
Dale O. Cloninger	−1.77	−2.23	63.79	97	11
Laurence H. Ross	−2.39	−2.10	67.74	63	7
Greg Pogarsky	−1.50	−2.03	53.24	61	6
Daniel S. Nagin	−0.73	−1.79	45.19	87	9
Theodore G. Chiricos, Gordon P. Waldo	−1.55	−1.61	49.51	81	8
David W. Rasmussen	−1.64	−1.45	28.13	81	8
Bruce L. Benson	−1.61	−1.45	29.94	92	9
Charles R. Tittle	−0.90	−1.38	47.71	61	6
Overall mean	−1.40	−1.37	41.66	6530	663
Raymond Paternoster	−1.55	−1.33	37.01	156	16
<u>Other</u>	−1.38	−1.29	40.69	5058	513
Horst Entorf	−2.51	−1.20	42.96	55	6
Ann D. Witte	−1.32	−1.20	38.21	66	7
Maynard L. Erickson	−1.22	−0.87	34.29	41	4
Jack P. Gibbs	−1.13	−0.85	31.19	61	6
Steven D. Levitt	−0.61	−0.83	32.00	112	11
Thomas B. Marvell	−1.17	−0.76	39.56	61	6
Alex R. Piquero	−0.84	−0.61	31.42	81	8
William C. Bailey	−0.35	−0.26	13.92	173	17

Mean and median correspond to the (normalized) t-values of the particular group. % is the percentage of estimates which are consistent with the deterrence hypothesis and significant at a 5% level in a two sided test. *#e* is the weighted number of all valid estimates. *#s* is the number of studies the estimates are based on. *Italic* entries are not significant in the ANOVA at a 5% level, but are included, like the overall mean, to show otherwise interesting or some selected groups. The <u>underlined</u> entry is the reference category. Gordon P. Waldo and Theodore G. Chiricos appear together in all of their studies.

As mentioned before, we have build an index which relates to the subjective quality of each study. It consists of the magnitude and quantity of problems reported by the author and the extent of unreported problems judged by the reader (i.e., the user who recorded the study). We have aggregated this index into three categories: good, medium and poor quality. Again, we emphasize that there is no such thing like a flawless study; not all problems can be coped with by corrective measures and some may lie deep in the available data source. While the estimates of poor studies differ significantly from those of medium quality, the order in table 3.37 is, all in all, somewhat inconclusive. Studies of medium quality (−1.25, 39.44%) are less in favor of the deterrence theory than those of poor (median −1.59, percentage 46.41%) and good quality (−1.65, 45.25%). Although the averaging effect of the sheer size of the category of medium quality may partially explain this, the order remains a bit strange. The first and last rows are practically identical (in regard to the mean (normalized) t-value) while their median and percentage differ.

Table 3.37: Differences by the quality of a study

Quality	mean	median	%	#e	#s
Good quality	−1.38	−1.65	45.25	1840	187
Poor quality	−1.77	−1.59	46.41	538	56
Overall mean	−1.40	−1.37	41.66	6530	663
Medium quality	−1.37	−1.25	39.44	4152	420

Mean and median correspond to the (normalized) t-values of the particular group. % is the percentage of estimates which are consistent with the deterrence hypothesis and significant at a 5% level in a two sided test. *#e* is the weighted number of all valid estimates. *#s* is the number of studies the estimates are based on. *Italic* entries are not significant in the ANOVA at a 5% level, but are included, like the overall mean, to show otherwise interesting or some selected groups. The underlined entry is the reference category.

Each study has to rely on some data. This can be public crime data (like the UCR in the USA, the PKS in Germany, public surveys or non-public data like the combination and linkage of various data sources, confidential data, experiments, self conducted surveys etc. Table 3.38 shows that there are no significant differences in the estimates when we categorize them according to the implemented data source. Nevertheless, it is worth mentioning that studies using the UCR come along with the smallest (normalized) t-values (−1.07, 36.99%), while those estimates based on non-public data yield "better" values (−1.68, 47.29%).

Table 3.38: Differences by the public data base

Data base	mean	median	%	#e	#s
PKS	−1.93	−1.78	37.32	75	8
None	−1.46	−1.68	47.29	1714	175
Overall mean	−1.38	−1.33	41.04	5900	600
Other	−1.33	−1.31	39.36	2618	264
UCR	−1.33	−1.07	36.99	1494	153

Mean and median correspond to the (normalized) t-values of the particular group. % is the percentage of estimates which are consistent with the deterrence hypothesis and significant at a 5% level in a two sided test. *#e* is the weighted number of all valid estimates. *#s* is the number of studies the estimates are based on. *Italic* entries are not significant in the ANOVA at a 5% level, but are included, like the overall mean, to show otherwise interesting or some selected groups. The underlined entry is the reference category.

As already described in subsection 3.3.3, we have an index of the general opinion of the author at our disposal, aggregated over all crimes. Table 3.39 is more or less a verification and plausibility check. As expected, all categories are significantly different from the reference category (partial approval) and all three statistics are in descending order (from −2.66, 64.5% to 0.32, 8.02%). Even the *Undefined* category is in the middle (−0.91, 35.41%), although it is not to be mixed up with something like "indifference"; it accumulates all not unambiguously definable opinions:

usually these happen to be studies lacking any usable statements which could reveal the opinion of the author or, which happens quite often, the opinion depends heavily on various conditions.

Table 3.39: Differences by the overall author opinion

Opinion	mean	median	%	#e	#s
Full approval	−3.07	−2.66	64.50	1377	142
Partial approval	−1.87	−2.08	54.48	2151	215
Overall mean	−1.40	−1.37	41.66	6530	663
Undefined	−0.75	−0.91	35.41	832	86
Partial disapproval	−0.58	−0.52	24.54	1158	118
Full disapproval	0.37	0.32	08.02	1012	102

Mean and median correspond to the (normalized) t-values of the particular group. % is the percentage of estimates which are consistent with the deterrence hypothesis and significant at a 5% level in a two sided test. *#e* is the weighted number of all valid estimates. *#s* is the number of studies the estimates are based on. *Italic* entries are not significant in the ANOVA at a 5% level, but are included, like the overall mean, to show otherwise interesting or some selected groups. The underlined entry is the reference category.

3.5.2 Estimate-Related Groups

Models are to be used, but not to be believed.

Henry Theil, Principles of Econometrics, 1971

Now, we will consider those grouping variables which may change for each estimate. Again, we present one table and a comment for each selected variable.

Although we have already classified the type of each study in table 3.35, we have another specification on the estimate-level. In table 3.40 we explicitly distinguish estimates which are concerned with the death penalty and aggregated survey, experiment and crime data. In contrast to the specification on the study-level, the types of each estimate can change within a study, although this is only rarely the case. The most striking fact of is at its very bottom: the estimates concerned with the death penalty (−0.43, 26.61%). Whether this depends on the large scepticism towards the deterrent effect[43] of the death penalty in the United States, or whether there is indeed no (measurable) effect, the table does not tell. The order of the other categories correspond to these of the specification on the study-level: experiments are associated with the most deterrent effects (−2.10, 55.59%), while crime data and surveys yield about the same results and are not significantly different.

Before looking at the individual offenses, table 3.41 reveals that estimates associated with the violation of rules in games (−2.01, 54.19%) or misdemeanors (−1.81, 49.23%) are significantly more often in accordance with the deterrence hypothesis. On the other hand, estimates associated with deviant behavior do support the deterrence hypothesis significantly less often (−0.97,

[43] In fact, only 5% of all estimates (from four studies) concerned with the death penalty, do not use U.S. data.

Table 3.40: Differences by the type of estimate

Type	mean	median	%	#e	#s
Experiment	−1.71	−2.10	55.59	832	85
Crime data	−1.48	−1.40	41.12	3569	387
Overall mean	−1.40	−1.37	41.66	6530	663
Survey	−1.35	−1.22	40.61	1595	168
Death Penalty	−0.57	−0.43	26.61	534	79

Mean and median correspond to the (normalized) t-values of the particular group. *%* is the percentage of estimates which are consistent with the deterrence hypothesis and significant at a 5% level in a two sided test. *#e* is the weighted number of all valid estimates. *#s* is the number of studies the estimates are based on. *Italic* entries are not significant in the ANOVA at a 5% level, but are included, like the overall mean, to show otherwise interesting or some selected groups. The underlined entry is the reference category.

41.99%). However, it should be noted that more than 80 percent of all estimates deal exclusively with crimes; only 8.5% are exclusively about non-crimes.

Table 3.41: Differences by the formal severity of an offense

Type	mean	median	%	#e	#s
Violating game-rules	−1.56	−2.01	54.19	158	16
Misdemeanors	−1.41	−1.81	49.23	620	76
Overall mean	−1.40	−1.37	41.66	6530	663
Crimes	−1.38	−1.34	41.25	5974	613
Other	−0.75	−1.34	34.39	96	11
Informal deviant behavior	−1.26	−0.97	41.99	113	19
Formal deviant behavior	−0.81	−0.60	29.34	161	23

Mean and median correspond to the (normalized) t-values of the particular group. *%* is the percentage of estimates which are consistent with the deterrence hypothesis and significant at a 5% level in a two sided test. *#e* is the weighted number of all valid estimates. *#s* is the number of studies the estimates are based on. *Italic* entries are not significant in the ANOVA at a 5% level, but are included, like the overall mean, to show otherwise interesting or some selected groups. The underlined entry is the reference category.

Researchers implement various variables to measure crime. The most common way is the usage of the respective crime rate for estimates based on crime data, and the self reported delinquency for surveys. While estimates using the crime rate are associated with less deterrent effects (−1.13, 37.63%) than the overall mean (−1.37, 41.79%), table 3.42 displays significantly "better" values (i.e., more supportive of the deterrence hypothesis) for estimates which use the absolute number of reported crimes (−1.90, 48.21%) or the probability of future delinquency (−2.03, 51.01%). Estimates using accidents (−1.69, 46.16%) or violations of prescriptive limits (−1.73, 46.39%) are also significantly "better". Some studies about general deterrence use recidivism[44] as as the

[44] Studies using recidivism are usually about specific deterrence. As mentioned in section 2.1, we have included some of these studies for various reasons.

exogenous variable. For these individuals there is no support for the deterrence hypothesis at all (0.53, 13.46%).

Table 3.42: Differences by the implemented endogenous variable

Endogenous variable	mean	median	%	#e	#s
Probability of future delinquency (surveyed is delinquent)	−1.38	−2.03	51.01	266	30
Number of reported crimes (absolute number)	−1.18	−1.90	48.21	722	77
Violating prescriptive limits	−1.70	−1.73	46.39	187	22
Other	−2.00	−1.71	46.97	668	88
Accidents	−1.43	−1.69	46.16	274	31
Self reported delinquency	−1.45	−1.45	43.44	1184	127
Overall mean	−1.41	−1.37	41.79	6429	657
Crime rate	−1.36	−1.13	37.63	2918	302
Probability of delinquency of fictitious offense (surveyed is delinquent)	−0.77	−0.94	27.43	78	10
Recidivism	0.23	0.53	13.46	44	5

Mean and median correspond to the (normalized) t-values of the particular group. % is the percentage of estimates which are consistent with the deterrence hypothesis and significant at a 5% level in a two sided test. *#e* is the weighted number of all valid estimates. *#s* is the number of studies the estimates are based on. *Italic* entries are not significant in the ANOVA at a 5% level, but are included, like the overall mean, to show otherwise interesting or some selected groups. The underlined entry is the reference category.

Table 3.43 shows the statistics distinguished by the various crime types. Very prominent is the upper part which is clearly dominated by non-violent crimes (malicious mischief being the only exception), while the lower part is exclusively made up by violent and drug-related crimes. The offenses which are in best accordance with the deterrence hypothesis are speeding (−2.21, 50.06%), tax evasion (−2.09, 53.04%), fraud (−1.90, 49.33%) and environmentally related offenses (−1.67, 46.57%). Sexual assault (−0.5, 28.67%; except rape, which is more in the middle), negligent assault (−0.36, 35.60%), manslaughter (0.04, 17.87%) and the possession of drugs are all only very weakly related to deterrent effects. Surprisingly, vehicle theft (−1.18, 39.48%), which is usually called the best property crime to measure deterrence (minimal reporting bias), is found somewhat below the overall mean. It is also noteworthy that dealing with soft drugs is more affected by deterrence than dealing with hard drugs (median of −0.63 and 0.04, percentage of 40.31% and 23.38%).

We also study the exogenous variables in each of the following categories: the death penalty, reported crimes, surveys or experiments. Table 3.44 shows the corresponding statistics of these four categories (the elements of each category are ordered by their median values). While the overall support of a deterrent effect of the death penalty is low, estimates using its existence have especially “bad” values (0.15, 11.97%), while the uncategorized estimates have very “good” values (−1.49, 39.96%). This may indicate that there are ways to detect some deterrent effect but

Table 3.43: Differences by the types of crime

Type	mean	median	%	#e	#s
Speeding	−1.53	−2.21	50.06	72	9
Tax evasion	−1.90	−2.09	53.04	474	53
Larceny (severe)	−1.42	−2.08	51.64	207	43
Drunk driving	−1.60	−2.00	50.72	787	92
Malicious mischief	−1.44	−2.00	50.80	93	19
Larceny (inferior)	−1.14	−2.00	51.78	173	35
Fraud	−1.72	−1.90	49.33	257	47
Other	−1.62	−1.68	43.60	460	60
Environmental crimes, Viol. of prescriptive limits	−0.99	−1.67	46.57	151	17
Burglary	−1.24	−1.44	43.91	795	182
Overall mean	−1.40	−1.37	41.55	6518	662
Larceny (Index I, general)	−1.30	−1.28	40.76	821	190
Other misdemeanors	−1.59	−1.27	43.73	206	29
Vehicle theft	−1.04	−1.18	39.48	558	133
Robbery	−1.28	−1.16	39.74	789	196
Rape	−1.44	−1.10	38.49	452	118
Other crimes	−1.09	−1.06	40.12	526	78
Homicide	−1.17	−0.88	34.41	1415	237
Crime rate (general)	−1.07	−0.87	32.81	402	64
Assault	−1.24	−0.81	38.07	661	167
Drug related crime (general)	−1.03	−0.72	40.13	288	55
Drug dealing (soft)	−1.22	−0.63	40.31	42	8
Sexual assault	−0.92	−0.50	28.67	33	17
Negligent assault	−0.85	−0.36	35.60	119	36
Manslaughter	−0.38	0.04	17.87	81	17
Drug dealing (hard)	−0.43	0.04	23.38	29	7

Mean and median correspond to the (normalized) t-values of the particular group. % is the percentage of estimates which are consistent with the deterrence hypothesis and significant at a 5% level in a two sided test. *#e* is the weighted number of all valid estimates. *#s* is the number of studies the estimates are based on. *Italic* entries are not significant in the ANOVA at a 5% level, but are included, like the overall mean, to show otherwise interesting or some selected groups. No reference category is used because multiple entries were common.

not with the usual measures (i.e., simple execution rates or law-dummies).

We have a much more detailed set of variables for reported crimes. Most deal with the probability of punishment and can be found in the upper part, while those related to the severity are in the lower part. As expected, the police expenditures and strength[45] have "bad" values (−1.01, 30.53% and 0.03, 21.67%). While most (except the incarceration rate) variables of the conviction cascade (see page 43) are more in favor of the deterrence hypothesis, rates with crime in the denominator have especially "good" values. Among the variables which measure the severity of

[45] By police strength we usually refer to all variables which measure the police force (e.g., the number of officers, employees, etc.).

punishment, the mean sentence length (−1.41, 33.64%, in opposition to the mean time in prison (−0.69, 22.08%) at the end of the table) and the regime shift dummy (−1.57, 39.90%) yield the "best" values.

Similar to the estimates based on reported crimes, those based on surveys can also be roughly divided into two parts: most variables which relate to the probability of detection and punishment are in the upper part of table 3.44, while those relating to the severity of punishment are in the lower part. In both cases, variables concerned with friends and family come before those dealing with justice. The estimates using the probability of punishment by friends or family have especially "good" values (−2.43, 57.66%).

Regarding experiments we see, again, basically the same picture. Using experimental and actual variation of the detection probability yield very "good" values (−3.05, 76.03% and −2.10, 57.87%), the estimates implementing the variation of the severity of punishment do the opposite (−1.16, 34.45% and −1.01, 41.51%). In both cases the experimental variation yield slightly "better" estimates than actual variation, indicating that deterrence can be more readily detected, when the parameters are more in control of the researcher.

Compared to the estimates from reported crimes, surveys and death penalties, the number of observations is most evenly spread for the experiment-categories. All in all, variables which correspond to probability measures are associated with "better" results than those dealing with the severity of punishment; this is true for all categories.

Table 3.44: Differences by the exogenous crime variable

Variable	mean	median	%	#e	#s
Death penalty					
Other	−1.19	−1.49	39.96	97	19
Percentage of all convictions	−0.72	−0.56	20.28	45	11
Execution rate	−0.65	−0.51	26.55	315	57
Overall mean	−0.63	−0.43	26.27	515	78
Existence of death penalty	0.17	0.15	11.97	67	14
Crime data					
Convicted per crime	−3.27	−3.46	81.99	41	11
Incarceration per crime	−2.60	−2.39	57.41	59	13
Probability dummy (regime shift)	−1.97	−2.15	63.93	228	31
Arrest rate	−2.10	−1.94	49.05	619	110
Clearance rate	−1.84	−1.93	47.41	344	65
Conviction rate	−2.30	−1.83	46.11	255	71
Fine	−1.95	−1.71	44.96	43	16
Other	−1.57	−1.66	45.09	558	101
Severity dummy (regime shift)	−1.73	−1.57	39.90	234	32
Mean sentence length (sentenced)	−1.20	−1.41	33.64	133	45
Overall mean	−1.47	−1.40	41.12	3550	386
Police expenditures	−0.39	−1.01	30.53	258	45

continued on the next page…

. . . last page of table 3.44 continued

Variable	mean	median	%	#e	#s
Probation rate	−1.17	−0.91	31.89	23	10
Inspections	−1.26	−0.82	31.84	79	11
Incarceration rate	−1.28	−0.77	30.10	42	18
Mean sentence length (served)	−0.68	−0.69	22.08	112	29
Incarcerations (absolute or per capita)	−0.80	−0.48	24.96	57	17
Police strength	−0.07	0.03	21.67	513	84
Surveys (all in expectations)					
Probability of punishment by friends or family	−2.05	−2.43	57.66	105	25
Probability of punishment by employment law	−1.76	−2.06	52.51	39	9
Probability of punishment by justice	−1.49	−1.68	45.83	292	53
Probability of detection by police	−1.47	−1.63	44.83	466	82
Overall mean	−1.36	−1.25	40.86	1563	168
Other	−2.27	−0.86	35.55	244	35
Severity of punishment by friends or family	−0.90	−0.86	23.73	62	15
Probability of punishment by others	−0.91	−0.83	24.31	59	13
Probability of other kind of punishment	−0.88	−0.80	12.18	26	5
Severity of punishment by justice	−0.82	−0.64	32.24	211	53
Previous experience with police or justice	−0.52	−0.40	31.82	56	6
Type of punishment	−0.46	−0.26	19.53	17	6
Severity of punishment by others	1.57	−0.07	32.35	29	6
Probability of detection by others	1.36	0.76	25.47	27	7
Experiments					
Experimental variation of probability of detection	−3.10	−3.05	76.03	136	19
Other	−1.85	−2.52	63.37	141	15
Actual variation of probability of detection	−1.35	−2.10	57.87	258	27
Overall mean	−1.75	−2.09	55.45	806	83
Experimental variation of severity of punishment	−1.12	−1.16	34.45	140	18
Actual variation of severity of punishment	−1.63	−1.01	41.51	137	19

Mean and median correspond to the (normalized) t-values of the particular group. % is the percentage of estimates which are consistent with the deterrence hypothesis and significant at a 5% level in a two sided test. *#e* is the weighted number of all valid estimates. *#s* is the number of studies the estimates are based on. *Italic* entries are not significant in the ANOVA at a 5% level, but are included, like the overall mean, to show otherwise interesting or some selected groups. The underlined entry is the reference category.

end of the table 3.44

It is well known that an omitted variable bias may pose a problem in studies about deterrence (see section 2.3.1). Depending on the included variables the significance of the estimates may vary considerably. As table 3.45 makes obvious, the inclusion of certain variables seem to make a relevant difference (the largest median is 0.04, the smallest −2.37). Naturally, some of the listed variables are only applicable for certain kinds of studies (e.g., GDP for studies analyzing nations or states, time trends for studies with a time dimension, etc.). Therefore, the listed categories may be strongly affected by other influences. Conditional on the usage of covariates, estimates which consider the GDP (−2.37, 51.85%), the labor force (−2.09, 52.49%), consumption (−2, 50.26%), drug usage (−1.94, 48.76%) and spatial fixed effects (−1.82, 48.76%) come along with values which are more in favor of the deterrence hypothesis. This could be evidence that the wealth of a nation and its consumption expenditures can be interpreted as proxies for the opportunity costs of crime. It also seems to be important to control for drug usage and unobserved heterogeneity, if applicable. By contrast, alcohol consumption (−0.44, 22.49%), social integration (−0.35, 22.73%), risk propensity and previous convictions (0.04, 17.8%) are accompanied with "bad" values. This could mean that these variables take over some of the effect of the implemented deterrence measures. The large difference between previous incarcerations and convictions (the median differs by 102%, the percentage by 63%) can be interpreted as a warning that such simple correlations only indicate relationships - neither do they imply cause, nor do they claim completeness. It is also noteworthy that income, unemployment, race, age, sex and youths are the most commonly used covariates. They are common variables in most data bases and are placed shortly under the overall mean.

Table 3.45: Differences by the used covariates

Variable	mean	median	%	#e	#s
GDP	−2.27	−2.37	51.85	89	12
Labor force	−2.23	−2.09	52.49	165	22
Consumption	−1.99	−2.00	50.26	130	16
Drug usage	−2.46	−1.94	48.76	84	11
Fixed effects (spatial)	−2.26	−1.82	48.74	651	80
Marital status	−2.36	−1.67	46.41	367	50
Previous incarceration	−1.10	−1.65	47.79	58	7
Time trend	−1.39	−1.63	36.71	351	44
Property value	−1.67	−1.61	42.24	123	15
Miles driven	−1.22	−1.59	36.85	85	11
Fixed effects (time)	−1.78	−1.44	40.16	782	88
Unemployment	−1.43	−1.40	40.04	1742	203
Overall mean	−1.41	−1.37	40.29	4682	523
Education	−1.77	−1.33	37.76	569	72
Youths	−1.60	−1.31	40.07	920	112
Other	−1.36	−1.31	39.12	4190	482
Income	−1.49	−1.28	38.02	1920	220
Sex	−1.47	−1.27	40.08	936	124
Race	−1.36	−1.17	38.49	1467	173
Age	−1.36	−1.07	39.13	1319	159
Morality	−1.12	−1.04	27.06	106	14
Random effects	−0.69	−1.01	18.87	66	10
Personal characteristics	−2.70	−1.00	43.22	193	23
Population (-growth)	−1.09	−0.98	33.81	745	92
Poverty, welfare	−0.54	−0.91	33.54	416	52

continued on the next page...

...last page of table 3.45 continued

Variable	mean	median	%	#e	#s
Acceptance of norms	−0.95	−0.88	33.56	140	23
Income inequality	−1.21	−0.84	31.67	522	64
Social class	−1.24	−0.80	28.66	52	7
Religion	−0.98	−0.77	30.11	112	16
Nationality	−0.97	−0.63	25.04	212	26
Urbanity	−0.77	−0.59	29.41	533	71
Alcohol (consumption)	−0.58	−0.44	22.49	109	16
Social integration	−0.51	−0.35	22.73	95	15
Risk propensity	−0.26	−0.25	14.22	50	7
Previous convictions	0.17	0.04	17.80	106	14

Mean and median correspond to the (normalized) t-values of the particular group. % is the percentage of estimates which are consistent with the deterrence hypothesis and significant at a 5% level in a two sided test. *#e* is the weighted number of all valid estimates. *#s* is the number of studies the estimates are based on. *Italic* entries are not significant in the ANOVA at a 5% level, but are included, like the overall mean, to show otherwise interesting or some selected groups. No reference category is used because multiple entries were common.

end of the table 3.45

In contrast to other meta analyses, our focus variables are not always the main variable in the included studies. Some authors use deterrence measures simply as covariates (e.g., analyzing unemployment and crime). In fact, there is a significant difference between the (normalized) t-values from estimates which focus on the deterrent effect and those which do not. The former have a mean value of −1.47 (median −1.44) while the latter values average to −1.02 (median −1.01). Moreover, the percentage of significant values, consistent with the deterrence hypothesis, falls from 42.80 to 35.04 (see table 3.46). However, there are at least two opposite explanations of this relationship. On the one hand, it could be that models in studies which incorporate deterrence variables as covariates, are rather inappropriate to measure deterrence (e.g., miss important variables or error corrections). Another explanation could be a publication bias: studies concentrating on deterrence are simply biased towards “better” results.

Table 3.46: Differences by the focus on deterrence

Focus	mean	median	%	#e	#s
<u>Main focus</u>	−1.47	−1.44	42.80	5564	575
Overall Mean	−1.40	−1.37	41.66	6530	663
Covariate	−1.02	−1.01	35.04	966	104

Mean and median correspond to the (normalized) t-values of the particular group. % is the percentage of estimates which are consistent with the deterrence hypothesis and significant at a 5% level in a two sided test. *#e* is the weighted number of all valid estimates. *#s* is the number of studies the estimates are based on. *Italic* entries are not significant in the ANOVA at a 5% level, but are included, like the overall mean, to show otherwise interesting or some selected groups. The <u>underlined</u> entry is the reference category.

Table 3.47 shows the statistics diversified by the implemented bivariate methods. These are obviously dominated by the (Pearson) correlations (38.7% of all applicable estimates), which are significantly closer to zero than all other (−0.68, 32.11%). Most other methods, even the (often plain and simple) differences of values (−2.08, 54.09%) yield results significantly more in favor of the deterrence hypothesis. It is obvious that the overall mean for bivariate methods seems to result mainly from the correlations - all other bivariate methods yield results which, more or less, strongly agree with the deterrence hypothesis. For the sake of completeness, we have also recorded whether there is any corrections for simultaneity (not applicable for bivariate methods) or other error-correction (no observation).

Table 3.47: Differences by the bivariate methods

Method	c	mean	median	%	#e	#s
Kendall's τ (ordinal)	no	−1.99	−2.67	57.26	56	7
Bivariate regression	no	−3.04	−2.66	73.84	32	9
Other	no	−1.99	−2.45	56.83	166	23
χ^2-test	no	−2.24	−2.39	68.37	97	12
Spearman's ρ (ordinal)	no	−1.88	−2.19	66.12	30	4
Differences (mean, percentages, etc.)	no	−1.67	−2.08	54.09	258	30
ANOVA	no	−1.59	−1.81	49.76	140	15
Point biserial correlation	no	−1.59	−1.37	49.82	25	5
Overall mean	no	−1.37	−1.37	45.13	1594	203
Gamma (ordinal)	no	−1.13	−1.16	36.42	148	20
Pearson correlation	no	−0.78	−0.68	32.11	617	95

Mean and median correspond to the (normalized) t-values of the particular group. *c* marks methods which somehow corrected for simultaneity or other problems. % is the percentage of estimates which are consistent with the deterrence hypothesis and significant at a 5% level in a two sided test. *#e* is the weighted number of all valid estimates. *#s* is the number of studies the estimates are based on. *Italic* entries are not significant in the ANOVA at a 5% level, but are included, like the overall mean, to show otherwise interesting or some selected groups. The underlined entry is the reference category.

The most frequent multivariate method is OLS (48.5%), followed by 2SLS (or more stages) with 14% and Logit or Probit models (10% of all multivariate methods). Since these methods are quite general, we recorded whether they have any corrections for simultaneity or implemented other error-corrections mechanisms. The results are given in table 3.48. We see that the estimates based on Tobit (−3.05, 66.12% and −1.89, 49.4%) or GLS-estimates (−2.8, 88.46% and −2.54, 63.1%) are significantly "better" than those based on simple OLS (−1.32, 40.26% and −1.46, 37.29%). Using partial correlations, as is the case with correlations in the bivariate case, yield the "worst" values (−0.1, 12.93%), on par with path analysis (0.22, 19.9%). However, some methods are restricted to certain disciplines; e.g., path analysis is not used by economists. There is no obvious tendency whether estimates based on methods with corrections do better correspond to the deterrence hypothesis or not.

Table 3.48: Differences by the multivariate methods

Method	c	mean	median	%	#e	#s
Tobit	yes	−2.55	−3.05	66.12	28	4
VAR	no	−3.06	−2.95	91.18	14	3
GLS	yes	−2.67	−2.80	88.46	26	4
GLS	no	−2.32	−2.54	63.10	83	11
ARIMA	no	−2.33	−2.42	63.07	67	8
Tobit	no	−2.02	−1.89	49.40	188	24
Poisson regression	no	−1.31	−1.76	39.74	19	5
2SLS, 3SLS	no	−1.40	−1.71	43.94	179	22
Poisson regression	yes	−1.18	−1.70	40.68	50	6
Logit, Probit	yes	−2.02	−1.68	23.76	40	5
ARIMA	yes	−1.49	−1.51	44.05	209	24
2SLS, 3SLS	yes	−1.40	−1.49	42.69	513	74
OLS	yes	−1.06	−1.46	37.29	578	85
Overall mean		−1.41	−1.37	40.53	4936	549
<u>OLS</u>	no	−1.33	−1.32	40.26	1816	245
Logit, Probit	no	−2.02	−1.01	40.62	461	53
VAR	yes	−2.13	−1.00	41.67	27	4
Other ML	yes	−1.22	−1.00	34.05	48	7
Other	yes	−1.05	−0.92	33.16	188	36
Other	no	−1.25	−0.56	31.36	281	41
Partial correlation	no	−0.66	−0.10	12.93	42	10
Path analysis	no	0.03	0.22	19.90	87	14

Mean and median correspond to the (normalized) t-values of the particular group. *c* marks methods which somehow corrected for simultaneity or other problems. % is the percentage of estimates which are consistent with the deterrence hypothesis and significant at a 5% level in a two sided test. *#e* is the weighted number of all valid estimates. *#s* is the number of studies the estimates are based on. *Italic* entries are not significant in the ANOVA at a 5% level, but are included, like the overall mean, to show otherwise interesting or some selected groups. The <u>underlined</u> entry is the reference category.

3.6 Multivariate Statistics

[. . .] when flawless analyses are not obtainable, the best alternative is to use several different approaches with the hope that the batch as a whole will give reliable information.

Eide et al. (1994)

We already saw many interesting potential relationships in section 3.5. These bivariate insights are valuable to get a feeling of the data and to pave the way for the tests of various hypotheses. However, whether these hypotheses are true or not, cannot be fully tested in such bivariate analyses, because the relevant variables are highly correlated with other variables. As Decker (1976) points out “the identification and measurement of the variables relevant to any scientific inquiry represents at once the most basic and primordial task”. Therefore, the identification of relevant variables will be the primary goal of this section and subsection 4.2.2.

The quote from Eide et al. (1994) fits this section perfectly. As a multitude of data and methodologies are used to study deterrence, we follow the same path in analyzing them. Since there is no unique and predefined way for such an analysis, we employ several methods and try to find trustworthy results.

3.6.1 The Variables

I pointed at the moon and some fool looked at my finger.

Zen saying

In principle, we have 306 unique variables at our disposal. After converting the nominal and ordinal variables into binary variables, we reach a maximum of 1603 variables (see section B.1 for the complete codebook). This conversion is necessary to include any arbitrary set of variables and to analyze them by various methods at any stage of an analysis. These converted measures include, beside others, binary variables for all authors, countries and journals which appear in the meta-data base.

Effectively, we have 6530 usable estimates in our meta-data base (henceforth called valid estimates). Using all variables would imply a ratio of at least 4 : 1 between the number of observations and variables, which is a considerably bad ratio. Furthermore, the number of variables is a limiting factor for methods which rely on brute force approaches, and the usage of more than 500 variables often prohibits such methods. Variables with almost no variation are also quite useless, since the explanatory power of a method is based on the variation of the variables; without sufficient variation, the variable's explanatory power quickly becomes a technical artifact. Last but not least, presenting variables with only little variation is of doubtful benefit and likely a waste of space. As already mentioned in subsection 3.2.3, we neglect those variables with less than seven (study-related) or 15 (estimate-related) observations.

Hence, for calculative and parsimony reasons, we gather all dummies from some selected nominal variables in a "other"-variable of its respective category. The limit is chosen manually and is usually located at a break in the data (the frequency falling by a large amount). These correspond to the following variables which are presented in section 3.3 in more detail.

- All authors who contribute at least six studies are incorporated as dummies. All other, with five or less studies, are gathered in a "other" dummy variable.

- Each country in which an author worked at the time of writing, which is represented at least six times, is included as a dummy variable. All other (four times or less) are summed in a single dummy.

- All countries subject to a study, with at least seven occurrences, are included as dummies. All other (four or less occurrences) are summed in a single dummy.

- All countries in which a study is published are assigned to a unique dummy, if it appears at least 18 times. All other (four or less occurrences) are gathered in a single dummy.

- All journals which appear at least seven times are included as dummies (all with five or less occurrences are gathered in a single variable).

These five steps lead to a total reduction to 523 variables. If not stated otherwise, all following analyses are based on these variables. The ratio between the number of variables and observations is now 12.5 : 1, which is still not good but at least acceptable. As before, the (normalized) t-values are used as endogenous variable. When interpreting the coefficients in the following analyses we have to keep in mind that the reference category of each included dummy may change, depending on the selected sets of variables.

3.6.2 Factor- and Cluster Analysis

I can prove anything by statistics except the truth.

George Canning

Confronted with many and strongly correlated variables, it is natural and common to try to condense these variables in some fashion. Factor- and cluster analysis are two well known analytical methods to accomplish this.

Factor Analysis

The main reason to perform a factor (or principal components) analysis is to come up with a much lower number of variables which still inherit most of the relevant information. In good cases these new variables can be reasonably interpreted as generalized influencing factors. This is often a good approach when dealing with a moderate number of variables which are, at least after a proper rotation, orthogonal to each other.

It is not surprising that a principal component analysis with all variables does not bring forth reasonable - or at least usable - results. Of 515 factors, 163 have an eigenvalue larger than one and 11 larger than five (the largest are 26.5, 12.6 and 10.5). Although the factors with the largest eigenvalues could, after applying different rotation methods, be interpreted in a reasonable way, they are not usable in the subsequent analysis[46]. Nevertheless, those entries with meaningful loadings always covered much less than a dozen variables and the factors are practically meaningless in predicting the (normalized) t-values (they did not perform much better than randomly chosen variables).

Cluster Analysis

Partitioning the data might be useful to condense or to improve the understanding of the variables. One problem is obviously the heterogeneity of the data structure and, consequently, the choice

[46] For example, the first factor contained information about survey characteristics, the third dealt with variables related to experiments, the fourth with reliability-variables, etc.

of a proper dissimilarity measure. Nonetheless, cluster analysis, regardless of the implemented method, seems to be unable to detect any relevant clusters.

For example, a single linkage cluster analysis produces one huge cluster containing most of the observations and many little clusters with only a few. Looking at the top 16 clusters, we get one containing 98.7% of all observations and 15 cluster with one to 18 observations. Other linkage algorithms result in similar structures.

So, instead of condensing variables or objects we select variables by their importance.

3.6.3 Ordinary Least Squares

In the literature, too much emphasis is put on statistical significance, implicitly assuming a statistically significant effect is economically meaningful in terms of size.

Florax and de Groot (2002)

Although (weighted) ordinary least squares is used in all of the following regression methods, we call those methods (simple) OLS which do not objectively select "important" variables. When possible, we use robust clustered standard errors (each study is treated as a cluster). Although the residuals of all estimates are significantly different from the corresponding normal distribution, a visual inspection of each plot reveals that the deviations do not seem to be severe. Since we will compare all implemented methods in section 4.2, we show the results of the OLS regressions of all variables (table 3.49) and the set of variables which are significant at a 10% level in the first regression (table 3.50). Variables which cause singularity problems are dropped by the algorithm (22 out of 515 variables). In all regressions each dummy is to be interpreted in comparison to its opposite property; e.g., the coefficient of the Author Isaac Ehrlich is to be compared to a study without the participation of Ehrlich. If both values of a dummy are included in a regression, they are compared to the missing values. In the following tables we include the variation of each variable. For non-metric variables it indicates the percentage of entries which differ from the most frequent entry. This means for dummy variables that the largest possible variation is fifty percent. This information is important when interpreting some variables which are almost constant (i.e., have a very low variation) and influence only very few estimates. For reasons of parsimony, we display only those variables in table 3.49 which are significant at a 25% level.

Table 3.49: Multivariate analysis - full OLS

Variable	var.	coef.	t	p
Study: not explorative	4.0	−0.546	−1.17	0.241
Study: measuring points	97.4	0.001	1.65	0.100
Study: year of first measure	84.4	−0.000	−1.70	0.089
Study: time span in months	63.2	−0.001	−1.44	0.150
Study: size of second population	0.9	0.001	1.18	0.238

continued on the next page...

. . . last page of table 3.49 continued

Variable	var.	coef.	t	p
Study: size of first realized sample	58.8	−0.000	−3.13	0.002
Study: size of second realized sample	2.4	−0.001	−3.12	0.002
Study: maximum age in first sample	5.7	−0.019	−1.82	0.069
Study: check for validity	2.1	4.587	1.24	0.214
Study: tests of significance	9.4	−0.604	−1.86	0.063
Study: number of bivariate estimates	32.4	0.013	1.98	0.048
Study: user, tr	48.6	−1.898	−1.95	0.052
Study: publication, journal article	13.7	0.545	1.50	0.135
Study: publication, working paper, report	4.9	−1.002	−1.22	0.224
Study: publication, not a dissertation or master thesis	2.3	−1.929	−2.32	0.021
Study: author, David W. Rasmussen	1.3	−2.134	−1.21	0.226
Study: author, Simon Hakim	1.0	−2.486	−2.11	0.035
Study: author, Raymond Paternoster	2.4	−1.388	−1.58	0.115
Study: author, Isaac Ehrlich	1.0	−1.294	−1.16	0.246
Study: author, Maynard I. Erickson	0.6	1.750	1.69	0.092
Study: author, Jack P. Gibbs	0.9	−1.982	−1.47	0.141
Study: author, Alex R. Piquero	1.3	−2.081	−1.95	0.052
Study: journal, Accident Analysis and Prevention	2.2	−1.634	−1.16	0.248
Study: journal, Studies on Alcohol	1.1	−1.665	−1.25	0.213
Study: author, Germany	3.2	2.641	1.51	0.132
Study: author, Switzerland	1.0	3.645	2.41	0.016
Study: author, Finland	0.8	−4.270	−2.33	0.020
Study: author, Australia	2.0	1.120	1.54	0.125
Study: author, Sweden	0.9	2.474	2.36	0.018
Study: author, other country	3.7	0.942	1.35	0.177
Study: author, criminology	11.3	0.907	1.77	0.077
Study: author, law	3.6	0.710	1.24	0.217
Study: publication, type not applicable	0.4	−3.226	−2.43	0.015
Study: institute, sociology	21.3	−0.723	−1.34	0.179
Study: experiment (laboratory)	4.4	−1.193	−1.21	0.228
Study: experiment (field, institutional initiative)	1.6	−1.746	−2.03	0.043
Study: first population, Canada	4.3	2.516	3.38	0.001
Study: first population, Netherlands	1.3	−1.594	−1.45	0.146
Study: sample base, first population, complete country	36.9	−1.624	−1.68	0.093
Study: sample base, first population, partial country	37.4	−2.006	−2.00	0.046
Study: sample base, second population, complete country	2.4	−2.811	−1.86	0.064
Study: sample base, second population, partial country	2.2	−4.687	−2.29	0.022
Study: sample unit, first population, miscellaneous	7.4	1.250	1.88	0.060
Study: sample unit, second population, individuals	1.7	2.792	1.61	0.108
Study: sample individuals, second population, population	2.6	5.075	2.47	0.014
Study: sample individuals, second population, miscellaneous	1.2	4.603	1.75	0.080
Study: complete sample	9.8	−0.773	−1.58	0.114
Study: PKS is public data base	1.2	2.496	2.30	0.022
Study: miscellaneous public data base	40.1	0.778	2.26	0.024

continued on the next page. . .

…last page of table 3.49 continued

Variable	var.	coef.	t	p
Study: UCR is public data base	22.9	0.506	1.32	0.187
Study: no public data base	26.3	0.702	2.17	0.030
Study: no class over-represented	0.9	1.840	1.41	0.158
Study: no disadvantaged group	0.6	−2.242	−2.12	0.035
Study: percentage of convicted > 75%	0.6	2.653	2.13	0.034
Study: main location > 500000 inhabitants	3.5	−1.869	−1.89	0.059
Study: main location < 5000 inhabitants	0.2	−7.691	−4.92	0.000
Study: does not claim to be representative	32.5	−0.301	−1.15	0.250
Study: claims to be representative	19.4	−0.547	−1.77	0.078
Study: does not check representativeness	26.9	0.395	1.71	0.087
Study: closed questions for pretest	21.5	1.374	2.45	0.014
Study: mixed questions for pretest	2.1	2.283	2.46	0.014
Study: Guttman reliability method	0.2	9.035	2.54	0.011
Study: miscellaneous reliability method	0.3	−3.552	−2.00	0.046
Study: correlational reliability method	0.2	−4.002	−2.03	0.043
Study: variables reliable	3.9	2.123	1.18	0.238
Study: validity test of some variables	1.5	−3.717	−1.25	0.213
Study: unknown if variables valid	0.3	−7.100	−1.58	0.114
Estimate: deterrence is focus-variable	14.8	−0.327	−1.18	0.240
Estimate: sub-sample	14.9	0.260	1.26	0.207
Estimate: sub-sample of males	1.4	−1.48	0.139	
Estimate: sub-sample of non-urban area	0.8	−1.753	−1.52	0.128
Estimate: exogenous, index mean	0.0	−2.071	−1.53	0.127
Estimate: exogenous, index items miscellaneous	0.2	1.671	1.17	0.242
Estimate: exogenous, index items standardized	0.2	2.550	1.58	0.115
Estimate: study type, death penalty	8.2	3.145	1.19	0.234
Estimate: exogenous, crime data, incarceration per crime	0.9	−1.316	−1.47	0.143
Estimate: exogenous, crime data, convicted per crime	0.6	−1.422	−1.43	0.152
Estimate: exogenous, survey, is no experiment	0.9	2.789	1.88	0.061
Estimate: exogenous, survey, probability of detection by police	7.1	−1.048	−2.98	0.003
Estimate: exogenous, survey, probability of punishment by justice	4.5	−1.235	−3.35	0.001
Estimate: exogenous, survey, severity of punishment by justice	3.2	−0.436	−1.24	0.217
Estimate: exogenous, survey, probability of other kind of punishment	0.4	−1.193	−1.39	0.166
Estimate: exogenous, survey, probability of detection by friends or family	0.2	−1.658	−2.23	0.026
Estimate: exogenous, survey, probability of punishment by friends or family	1.6	−1.634	−3.97	0.000
Estimate: exogenous, survey, severity of punishment by friends or family	1.0	−0.678	−1.49	0.136
Estimate: exogenous, survey, time between offense and clearance	0.1	−1.293	−2.03	0.043
Estimate: exogenous, survey, relates to the present	21.3	2.299	1.39	0.164
Estimate: exogenous, survey, relates to the past	2.7	2.723	1.61	0.107
Estimate: exogenous, experiment, yes	7.2	3.426	1.95	0.052
Estimate: exogenous, experiment, no	5.4	3.397	1.89	0.060
Estimate: exogenous, experiment, relates to the present	13.9	−0.841	−1.47	0.142
Estimate: exogenous, experiment, relates to the past	0.5	−1.758	−1.32	0.188
Estimate: exogenous, relates to one year	42.7	0.663	1.62	0.107

continued on the next page…

. . . last page of table 3.49 continued

Variable	var.	coef.	t	p
Estimate: exogenous, relates to more than one year	12.1	−0.492	−1.20	0.232
Estimate: exogenous, metric category	36.8	1.959	1.80	0.073
Estimate: exogenous, interval category	9.1	1.746	1.46	0.145
Estimate: exogenous, binary category	18.8	1.923	1.75	0.080
Estimate: exogenous, nominal category	0.3	−3.318	−1.44	0.149
Estimate: exogenous, ordinal category	7.4	2.475	2.16	0.031
Estimate: exogenous, in differences	3.1	1.533	1.89	0.059
Estimate: endogenous, index miscellaneous	0.1	1.792	1.18	0.240
Estimate: endogenous, index additive, weighted	0.1	−1.001	−1.53	0.127
Estimate: endogenous, number of registered suspects	0.6	1.601	1.33	0.183
Estimate: endogenous, number of convicted to prison sentence	0.2	2.476	1.98	0.048
Estimate: endogenous, probability of delinquency of fictitious offense (surveyed is delinquent)	1.2	1.895	1.77	0.077
Estimate: endogenous, recidivism	0.7	2.768	2.28	0.023
Estimate: endogenous, accidents	4.2	1.232	1.49	0.136
Estimate: endogenous, self reported delinquency since age of fourteen	0.6	2.619	2.73	0.007
Estimate: crime category, misdemeanors	9.5	−1.137	−3.04	0.002
Estimate: crime category, formal deviant behavior	2.5	0.735	1.37	0.170
Estimate: crime category, other	1.5	1.042	1.44	0.150
Estimate: offense, assault	10.1	−0.297	−1.24	0.214
Estimate: offense, negligent assault	1.8	0.695	1.74	0.083
Estimate: offense, burglary	12.2	0.251	1.40	0.163
Estimate: offense, larceny (severe)	3.2	−0.591	−1.89	0.059
Estimate: offense, drug possession (hard)	0.5	−2.585	−1.87	0.061
Estimate: offense, driving without a licence	0.0	−0.858	−1.38	0.168
Estimate: offense, drunk driving	12.1	0.638	1.76	0.079
Estimate: offense, fare dodging	0.4	0.592	1.51	0.132
Estimate: offense, fraud	3.9	0.531	1.55	0.121
Estimate: offense, tax evasion	7.3	0.600	1.54	0.124
Estimate: offense, other	7.1	−0.721	−1.86	0.064
Estimate: offense, vehicle theft	8.5	0.267	1.36	0.175
Estimate: offense, environmental crimes, violations of prescriptive limits	2.3	1.606	1.61	0.107
Estimate: property and violent characteristics	48.8	0.627	1.41	0.159
Estimate: endogenous, metric category	19.5	−1.749	−1.60	0.111
Estimate: endogenous, interval category	4.1	−1.532	−1.26	0.208
Estimate: endogenous, ordinal category	4.6	−3.203	−2.85	0.004
Estimate: endogenous, binary category	9.6	−3.269	−2.92	0.004
Estimate: endogenous, not in logs	32.0	1.232	1.73	0.085
Estimate: endogenous, in logs	26.8	1.285	1.74	0.083
Estimate: endogenous, other transformation	7.8	−2.009	−2.18	0.030
Estimate: covariate, age	20.2	−0.491	−2.02	0.044
Estimate: covariate, marital status	5.6	0.701	1.88	0.060
Estimate: covariate, profession	0.4	1.677	1.89	0.060
Estimate: covariate, social class	1.5	−1.489	−1.72	0.086

continued on the next page. . .

. . . last page of table 3.49 continued

Variable	var.	coef.	t	p
Estimate: covariate, drug usage	1.3	−1.124	−1.26	0.209
Estimate: covariate, morality	1.6	1.020	2.02	0.044
Estimate: covariate, personal characteristics	3.0	−0.720	−1.54	0.124
Estimate: covariate, random effects	1.0	1.054	1.45	0.147
Estimate: covariate, poverty, welfare	6.4	0.673	2.34	0.020
Estimate: covariate, urbanity	8.2	0.446	1.48	0.138
Estimate: covariate, GDP	1.4	−1.005	−1.53	0.127
Estimate: covariate, population (-growth)	11.5	0.434	1.64	0.102
Estimate: covariate, alcohol (consumption)	1.7	0.623	1.32	0.188
Estimate: covariate, consumption	2.0	−0.717	−1.41	0.159
Estimate: covariate, risk propensity	0.8	1.905	2.63	0.009
Estimate: no correction for simultaneity	19.3	−1.357	−1.98	0.048
Estimate: unweighted model	6.8	−0.477	−1.25	0.211
Estimate: bivariate method, ρ	0.5	−2.774	−1.57	0.117
Estimate: bivariate method, binomial	0.2	2.063	1.26	0.207
Estimate: multivariate method, COX regression	0.3	2.331	1.17	0.242
Estimate: square root of sample size for negative values	79.2	−0.014	−4.44	0.000
Estimate: square root of sample size for positive values	82.5	0.052	6.90	0.000

$N = 6530$, $R^2 = 0.478$, number of cluster is 663, 22 out of 515 variables are dropped due to singularity problems. The column *var* refers to the variation of a variable (i.e., the percentage of valid observations); the maximum variation for dummy variables is fifty percent. c and t are the coefficients and the corresponding (normalized) t-values of the included variables. The reference category for dummies is usually the opposite or, in the case of multiple categories, the missing values.

end of the table 3.49

Since we use the set of variables which are significant at a 10% level as the most simple type of variable selection in section 4.2, we present those results in table 3.50.

Table 3.50: Multivariate analysis - OLS of 10%-significant variables

Variable	var.	coef.	t	p
Study: measuring points	97.4	−0.000	−1.54	0.123
Study: year of first measure	84.4	0.000	1.33	0.184
Study: size of first realized sample	58.8	−0.000	−4.99	0.000
Study: size of second realized sample	2.4	−0.001	−7.21	0.000
Study: maximum age in first sample	5.7	−0.006	−1.14	0.256
Study: tests of significance	9.4	−0.661	−2.77	0.006
Study: number of bivariate estimates	32.4	0.009	1.82	0.070
Study: user, tr	48.6	−0.783	−3.78	0.000
Study: publication, not dissertation or master thesis	2.3	0.163	0.42	0.673
Study: author, Simon Hakim	1.0	−0.534	−0.63	0.527
Study: author, Maynard I. Erickson	0.6	0.451	1.23	0.218

continued on the next page. . .

. . . last page of table 3.50 continued

Variable	var.	coef.	t	p
Study: author, Alex R. Piquero	1.3	0.181	0.57	0.566
Study: author, Switzerland	1.0	0.566	1.62	0.106
Study: author, Finland	0.8	−1.732	−4.95	0.000
Study: author, Sweden	0.9	0.950	1.82	0.069
Study: author, criminology	11.3	0.267	1.03	0.304
Study: publication, type not applicable	0.4	−1.971	−2.43	0.015
Study: experiment (field, institutional initiative)	1.6	−0.877	−1.24	0.215
Study: first population, Canada	4.3	0.769	3.62	0.000
Study: sample base, first population, complete country	36.9	−1.694	−1.14	0.255
Study: sample base, first population, partial country	37.4	−1.896	−1.28	0.200
Study: sample base, second population, complete country	2.4	−1.637	−2.71	0.007
Study: sample base, second population, partial country	2.2	−1.306	−1.53	0.127
Study: sample unit, first population, miscellaneous	7.4	0.270	0.80	0.423
Study: sample individuals, second population, population	2.6	2.257	2.89	0.004
Study: sample individuals, second population, miscellaneous	1.2	1.920	2.42	0.016
Study: PKS is public data base	1.2	0.323	0.95	0.341
Study: miscellaneous public data base	40.1	0.433	2.20	0.028
Study: no public data base	26.3	0.316	1.48	0.139
Study: no disadvantaged group	0.6	−0.734	−1.88	0.061
Study: percentage of convicted $> 75\%$	0.6	0.862	2.42	0.016
Study: main location > 500000 inhabitants	3.5	−1.583	−2.01	0.045
Study: main location < 5000 inhabitants	0.2	−2.453	−4.21	0.000
Study: claims to be representative	19.4	−0.140	−0.79	0.430
Study: does not check representativeness	26.9	0.478	3.01	0.003
Study: closed questions for pretest	21.5	1.100	4.15	0.000
Study: mixed questions for pretest	2.1	1.666	4.81	0.000
Study: Guttman reliability method	0.2	1.097	2.66	0.008
Study: miscellaneous reliability method	0.3	−2.030	−1.35	0.176
Study: correlational reliability method	0.2	−1.476	−4.96	0.000
Estimate: exogenous, survey, is no experiment	0.9	1.763	2.94	0.003
Estimate: exogenous, survey, probability of detection by police	7.1	−0.675	−3.26	0.001
Estimate: exogenous, survey, probability of punishment by justice	4.5	−0.654	−3.10	0.002
Estimate: exogenous, survey, probability of detection by friends or family	0.2	−1.330	−2.52	0.012
Estimate: exogenous, survey, probability of punishment by friends or family	1.6	−1.291	−4.52	0.000
Estimate: exogenous, survey, time between offense and clearance	0.1	−1.467	−2.88	0.004
Estimate: exogenous, experiment, yes	7.2	−0.688	−1.95	0.051
Estimate: exogenous, experiment, no	5.4	0.282	0.72	0.469
Estimate: exogenous, metric category	36.8	0.592	2.50	0.013
Estimate: exogenous, binary category	18.8	0.367	1.41	0.160
Estimate: exogenous, ordinal category	7.4	0.119	0.45	0.652
Estimate: exogenous, in differences	3.1	1.408	2.63	0.009
Estimate: endogenous, number of convicted to prison sentence	0.2	1.348	2.20	0.028
Estimate: endogenous, probability of delinquency of fictitious offense (surveyed is delinquent)	1.2	0.021	0.04	0.965

continued on the next page. . .

. . . last page of table 3.50 continued

Variable	var.	coef.	t	p
Estimate: endogenous, recidivism	0.7	1.730	3.63	0.000
Estimate: endogenous, self reported delinquency since age of fourteen	0.6	0.692	1.50	0.135
Estimate: crime category, misdemeanors	9.5	−0.111	−0.46	0.642
Estimate: offense, negligent assault	1.8	0.702	1.75	0.080
Estimate: offense, larceny (severe)	3.2	0.002	0.01	0.995
Estimate: offense, drug possession (hard)	0.5	−1.956	−1.75	0.080
Estimate: offense, drunk driving	12.1	−0.144	−0.66	0.507
Estimate: offense, other	7.1	−0.191	−0.74	0.461
Estimate: endogenous, ordinal category	4.6	−0.565	−1.99	0.047
Estimate: endogenous, binary category	9.6	−0.681	−2.60	0.009
Estimate: endogenous, not in logs	32.0	0.164	0.51	0.611
Estimate: endogenous, in logs	26.8	−0.516	−1.40	0.163
Estimate: endogenous, other transformation	7.8	−0.904	−1.57	0.118
Estimate: covariate, age	20.2	−0.067	−0.34	0.735
Estimate: covariate, marital status	5.6	−0.028	−0.10	0.920
Estimate: covariate, profession	0.4	0.090	0.15	0.882
Estimate: covariate, social class	1.5	−0.620	−0.67	0.505
Estimate: covariate, morality	1.6	0.019	0.07	0.940
Estimate: covariate, poverty, welfare	6.4	1.124	3.50	0.000
Estimate: covariate, risk propensity	0.8	0.664	1.79	0.073
Estimate: no correction for simultaneity	19.3	−0.539	−2.36	0.019
Estimate: square root of sample size for negative values	79.2	−0.013	−3.58	0.000
Estimate: square root of sample size for positive values	82.5	0.054	6.85	0.000

$N = 6530$, $R^2 = 0.304$, number of cluster is 663, all variables from table 3.49 which were significant at a 10% level are selected. The column *var* refers to the variation of a variable (i.e., the percentage of valid observations); the maximum variation for dummy variables is fifty percent. c and t are the coefficients and the corresponding (normalized) t-values of the included variables. The reference category for dummies is usually the opposite property or, in the case of multiple categories, the missing values.

end of the table 3.50

While most significant variables remain unchanged in terms of size and significance when proceeding from the large set to the smaller set, some variables change. Being not a dissertation or master thesis (which applies to almost all studies) changes from being significant and negative to positive insignificance. The indicator for Alex R. Piquero reverses its sign while the signs of all other authors remain unchanged. This could be explainable when studies from that author have some special properties which are not taken into account in the second regression. Severe larceny switches from negative significance to positive insignificance, while drunk driving does exactly the opposite, as well as the dummy indicating the logarithm of the endogenous variable. Finally, the impact of most covariates is largely reduced in significance.

All in all, important factors correlated with support of the deterrence hypothesis are the economic background in general (represented by the user tr who was responsible for all economic studies), Finnish studies, very large or small locations, studies which check the reliability of variables with correlations, use the probability and severity of punishment (and the celerity) by officials or friends and family in surveys, as well as estimates which are not corrected for simultaneity. The opposite effect can be found when Canadian data is studied, when "other" public data bases are used, when the studied individuals have almost all been convicted before, when authors do not check representativeness, when closed or mixed questions are used in a pretest, when the exogenous variables is metric or measured in differences, when the deterrence variable relates to prison sentences or recidivism, and, finally, covariates relating to poverty and welfare are implemented. Last but not least, the technical influence of the sample size and the size of the studied population (which strongly correlates with the sample size) have to be mentioned.

When the results are compared with the bivariate analysis in section 3.5, noteworthy changes are: German authors, when controlling for other effects, are now correlated with less support of the deterrence theory, while studies from Alex R. Piquero are now associated with more support in the larger set (table 3.49). The impact of many variables, which measure deterrence in surveys and appeared to be associated with less support in section 3.5 is now reversed. The coefficients of the covariates age, marital status and the social class switch their signs. Curiously, the correlation between the studied offenses and the resulting (normalized) t-values are rather incompatible with those from table 3.43.

3.6.4 Extreme Bounds Analysis

It would appear that we have reached the limits of what is possible to achieve with computer technology, although one should be careful with such statements, as they tend to sound pretty silly in 5 years.

John von Neumann, 1949

The principle of EBA is to regress all possible combination of k (out of N) exogenous variables on the variable to be explained, and to track the distribution of the associated t-values, refer also toe Leamer (1983, 1985) or Levine and Renelt (1992). After that, conclusions are derived from analyzing the distribution of these t-values. In every regression we include an identical set of five variables which we previously classified to belong to an appropriate model: the year of publication, if it is published in a journal, whether the user tr recorded the study and, for reasons explained in section 3.4, the square root of the sample size for negative and positive t-values. For each regression we record the coefficient, the t-value, the R^2 and the number of included observations. Inspired by Smith and Huang (1995), and in accordance to our previous analysis, we use clustered standard deviations in all regressions.

After doing the calculations the main task is to identify the most important and reliable variables. To do this, we use three kinds of EBA-test criteria (a variation of those found in Florax and de Groot (2002)) which are based on the statistical distribution of the tracked t-values:

- A Cumulative Distribution Function (CDF) Test (A): a variable is considered important, if the 1% and 99% quantiles share the same sign.

- A strong sign test (B): the influence of a variable is considered important, if all of its t-values are of the same sign.

- An extreme CDF-Test (C): a variable is considered important, if the 99% confidence intervals around the minimum and maximum of a variable do not include zero.

- To compare EBA with other methods which select only a few variables (about 50 variables), we add an absurd test (D), which includes a variable only if a $(1-7.342\cdot 10^{-51})\%$ confidence interval around the mean does not include zero.

Since every variable is included in at least 128777 regressions[47], there are still several variables which pass test D easily. There is no specific ordering of the tests except that any variable which passes test B also passed test A. Even a variable which passes the most restrictive test D may not pass test C if the distribution of the variable's t-values has rather long tails. However, we should not be too impressed by the results, reminding the advice given by Lovell (1983): "it is ironic that the data mining procedure that is most likely to produce regression results that appear impressive in terms of the customary criteria is also likely to be the most misleading in terms of what it asserts about the underlying process generating the data under study."

However, we do not use EBA to find any underlying model structure but to find variables which seem to have a stable influence in regard to the results a study provides. For reasons of robustness, we may not want to include variables with very asymmetrical tails, which might indicate an unstable distribution. Therefore, by calculating two stability coefficients, we exclude those variables X, which do not satisfy either

$$(X_{99\%}-\overline{X})/(\overline{X}-X_{1\%})\in[0.75,1.25] \text{ or } (\overline{X}/X_{50\%})\in[0.99,1.01],$$

whereas $X_{y\%}$ is the $y\%$-quantile of the distribution of variable X.

Although EBA is often used in economics (e.g., Levine and Renelt (1992); Sala-I-Martin (1997) or Bartley et al. (1998)) and also in deterrence studies (e.g., McManus (1985) or McAleer and Veall (1989)), it has several disadvantages. First, computing the statistics of all $\binom{N}{k}$ combinations is, in practice, computationally impossible for even small k if N is large. Our own implementation in STATA (Statacorp LP, 2005) requires one gigabyte of data per 3.6 million regressions and has a runtime of $O(N^k)$. Therefore, the largest k possible is three, resulting in $\binom{515-5}{3}=21978620$ regressions generating about 6GB of data, requiring about one week of calculation on a 4.2GHz Athlon X2 using only one core ($k=4$ would result in 2785 million regressions, taking 127x more time to run and would generate about 1TB of data; it is unlikely that an optimized version of the algorithm would make this feasible). Second, the vote-counting problem (Hedges and Olkin, 1985)

[47] This is the number of observations on which the interpretation of each variable is based.

might increase the difficulty of interpreting the distributions, because errors of the second type increase with the number of observations (in this case combinations). Third, it is an unresolved problem whether the calculated t-values should be weighted, e.g., with the ML of the respective model (Sala-I-Martin, 1997). Weighting can improve the quality of the conclusions (minimizing the influence of obviously improper models) or dampen them (since all models, whether they have a high ML or not, will most likely suffer severely from omitted variable bias).

Results from EBA could be further analyzed by applying a Response Surface Analysis (RSA), as sketched by Florax and de Groot (2002). However, this is not done here due to computational constraints (in combination with an EBA, the algorithm has a runtime of $O(N^{k+1})$). Instead, we apply a simple check for tail characteristics using the stability coefficients described above.

While interpreting the results, displayed in table 3.51, we have to keep in mind that all variables are accompanied by seven other variables of which five appear in every regression. Thus, the linear effects of the year of publication, the journal as media of publication, the overall influence of the user tr (i.e., the study was recorded in Darmstadt) and the number of observations are already taken care of. The reported statistics refer to the t-values across the regressions of all three-dimensional (unordered) subsets of the remaining variables. Negative values indicate that a variable enhances the compliance with the deterrence hypothesis (i.e., the null hypothesis is discarded with larger confidence) as it increases (e.g., when the dummy variable is one). Positive values indicate the reverse, while values near zero imply that a variable is not a significant determinant in a study about general deterrence. We have to keep in mind that a variable, which is almost never significant (e.g., most of its absolute t-values are below 1.96), but is negative (or positive) in all cases, still indicates that it may have a strong influence on the (normalized) t-values. For reasons of parsimony, we only display those variables which pass at least two tests (244 variables are not shown).

Table 3.51: Multivariate analysis - extreme bounds analysis

Variable	var	A	B	C	D	mean	sd	min	1%	50%	99%	max
Study: publication, number	83.3	1	1	0	0	1.70	0.17	0.22	1.10	1.72	2.18	3.07
Study: publication, year	100.0	1	1	0	0	−1.46	0.14	−2.49	−1.87	−1.46	−1.07	−0.31
Study: measuring points	97.4	1	1	1	0	−1.50	0.13	−2.36	−1.88	−1.50	−1.06	−0.44
Study: size of first realized sample	58.8	1	1	1	1	−4.10	0.06	−4.64	−4.31	−4.10	−3.91	−3.63
Study: rate of return of second sample	0.9	1	1	1	1	−1.61	0.08	−2.36	−1.93	−1.61	−1.31	−0.45
Study: error and plausibility checks	31.4	1	1	1	0	−0.80	0.11	−1.46	−1.13	−0.78	−0.56	−0.30
Study: tests of significance	9.4	1	1	1	1	−2.21	0.13	−3.58	−2.55	−2.21	−1.87	−1.54
Study: number of bivariate estimates	32.4	1	1	1	0	1.86	0.14	0.83	1.40	1.85	2.34	2.90
Study: user, mw	1.4	1	1	1	1	1.28	0.07	0.92	1.10	1.27	1.55	1.83
Study: publication, miscellaneous type	3.3	1	1	1	0	0.78	0.07	0.30	0.58	0.78	0.95	1.15
Study: publication, book	1.5	1	1	1	0	−0.52	0.06	−0.87	−0.68	−0.52	−0.35	−0.17
Study: author, Steven D. Levitt	1.7	1	1	1	1	1.63	0.09	0.95	1.41	1.62	1.99	2.52
Study: author, Simon Hakim	1.0	1	1	1	1	−1.35	0.07	−2.26	−1.52	−1.35	−1.17	−0.91

continued on the next page...

. . . last page of table 3.51 continued

Variable	var	A	B	C	D	mean	sd	min	1%	50%	99%	max
Study: author, Daniel S. Nagin	1.3	1	1	1	1	1.30	0.08	0.52	1.01	1.29	1.55	1.80
Study: author, Raymond Paternoster	2.4	1	1	0	0	−1.31	0.17	−2.50	−1.76	−1.33	−0.80	−0.09
Study: author, Isaac Ehrlich	1.0	1	1	1	1	−3.73	0.21	−5.43	−4.19	−3.74	−2.85	−1.80
Study: author, Harold G. Grasmick	1.6	1	1	1	0	−2.15	0.17	−4.94	−2.52	−2.16	−1.80	−1.45
Study: author, Laurence H. Ross	1.0	1	1	0	0	−0.79	0.06	−1.34	−0.94	−0.80	−0.59	−0.05
Study: journal, Economic Inquiry	1.4	1	1	1	1	−1.23	0.04	−2.02	−1.38	−1.23	−1.10	−0.68
Study: journal, Criminal Justice	2.2	1	1	1	1	2.53	0.12	1.25	2.23	2.53	2.88	3.15
Study: journal, Law and Economics	2.8	1	1	1	1	−1.58	0.08	−2.13	−1.79	−1.57	−1.35	−0.86
Study: journal, American Economic Review	2.0	1	1	0	0	−1.66	0.15	−3.09	−2.19	−1.66	−1.25	−0.35
Study: journal, Social Forces	2.2	1	1	1	0	−2.39	0.18	−3.75	−2.91	−2.39	−1.78	−0.80
Study: journal, American Journal of Economics and Sociology	1.7	1	1	1	1	0.96	0.06	0.43	0.74	0.96	1.10	1.52
Study: journal, Review of Economics and Statistics	1.6	1	1	1	1	−1.12	0.06	−1.45	−1.27	−1.13	−0.92	−0.53
Study: journal, Crime and Delinquency	1.1	1	1	0	0	2.01	0.17	0.37	1.52	2.01	2.65	3.80
Study: journal, Quantitative Criminology	1.4	1	1	0	1	0.89	0.06	0.14	0.71	0.90	1.01	1.24
Study: journal, Southern Economic Journal	1.3	1	1	0	0	1.57	0.14	0.29	1.21	1.58	1.98	2.91
Study: journal, Social Science Quarterly	2.2	1	1	1	0	−2.22	0.16	−3.29	−2.65	−2.21	−1.74	−0.49
Study: journal, Legal Studies	2.0	1	1	1	1	−1.24	0.08	−1.82	−1.51	−1.24	−0.98	−0.66
Study: publication, USA	21.9	1	1	1	1	−1.79	0.11	−2.49	−2.12	−1.79	−1.47	−0.70
Study: publication, Netherlands	2.6	1	1	1	1	2.55	0.09	1.73	2.34	2.54	2.85	3.42
Study: author, Canada	5.2	1	1	0	1	3.09	0.20	0.47	2.41	3.11	3.42	3.82
Study: author, Switzerland	1.0	1	1	1	1	3.04	0.16	2.25	2.69	3.01	3.70	4.53
Study: author, Sweden	0.9	1	1	1	1	2.48	0.11	1.58	2.17	2.49	2.78	3.29
Study: author, psychology	4.2	1	1	1	1	−3.31	0.16	−4.24	−3.59	−3.34	−2.61	−1.66
Study: publication, economics	36.0	1	1	1	1	−1.80	0.11	−2.47	−2.07	−1.80	−1.37	−0.64
Study: publication, criminology	20.4	1	1	1	1	3.92	0.12	2.93	3.51	3.92	4.21	4.61
Study: institute, law	3.9	1	1	1	0	−1.04	0.08	−1.68	−1.24	−1.04	−0.78	−0.38
Study: institute, miscellaneous	14.9	1	1	1	1	3.25	0.13	1.94	2.86	3.24	3.65	4.12
Study: cross section	23.2	1	1	1	0	1.61	0.16	0.68	1.21	1.62	2.36	3.14
Study: experiment (field, institutional initiative)	1.6	1	1	1	1	−1.65	0.09	−2.06	−1.83	−1.66	−1.19	−0.80
Study: first population, United Kingdom	5.0	1	1	1	1	−1.29	0.07	−1.93	−1.57	−1.29	−1.13	−0.93
Study: first population, Canada	4.3	1	1	1	1	3.87	0.21	1.35	2.91	3.90	4.28	5.10
Study: first population, Finland	0.9	1	1	0	0	−4.97	0.61	−12.25	−6.31	−4.93	−3.72	−0.54
Study: first population, Australia	1.7	1	1	0	0	−1.06	0.12	−2.67	−1.56	−1.05	−0.75	−0.05
Study: first population, other country	6.2	1	1	0	0	0.79	0.08	0.22	0.57	0.78	1.03	1.90
Study: sample base, second population, complete country	2.4	1	1	0	1	−2.05	0.13	−2.88	−2.29	−2.07	−1.43	−0.30
Study: sample unit, first population, states	21.9	1	1	1	1	2.80	0.13	2.10	2.43	2.81	3.22	3.62
Study: sample individuals, second population, students	0.5	1	1	1	0	−3.41	0.25	−7.68	−3.99	−3.44	−2.42	−1.16
Study: sample of extreme groups	0.5	1	1	0	0	1.37	0.16	0.14	0.97	1.36	1.96	3.72
Study: complete sample	9.8	1	1	1	1	−1.21	0.06	−1.63	−1.37	−1.21	−1.06	−0.85
Study: miscellaneous public data base	40.1	1	1	1	0	1.26	0.13	0.37	0.83	1.28	1.62	2.55
Study: income above average	0.6	1	1	0	0	−0.80	0.09	−1.94	−1.13	−0.80	−0.57	−0.16
Study: upper class over-represented	0.6	1	1	0	0	−0.80	0.09	−1.94	−1.13	−0.80	−0.57	−0.16
Study: main location > 500000 inhabitants	3.5	1	1	1	1	−1.32	0.05	−1.72	−1.45	−1.32	−1.18	−0.99
Study: main location 100000-500000 inhabitants	1.4	1	1	0	0	0.33	0.05	0.02	0.20	0.32	0.49	0.86

continued on the next page. . .

. . . last page of table 3.51 continued

Variable	var	A	B	C	D	mean	sd	min	1%	50%	99%	max
Study: mixed location	0.5	1	1	0	0	0.55	0.06	0.03	0.38	0.55	0.72	1.14
Study: does not check representativeness	26.9	1	1	1	1	2.46	0.12	1.73	2.15	2.44	2.83	3.32
Study: conditions for significance check fulfilled	49.2	1	1	0	0	0.80	0.09	0.09	0.54	0.80	1.06	1.46
Estimate: number of used covariates	74.1	1	1	0	0	1.03	0.10	0.21	0.68	1.04	1.29	1.67
Estimate: deterrence is focus-variable	14.8	1	1	1	1	−1.38	0.08	−1.99	−1.63	−1.39	−1.16	−0.63
Estimate: deterrence is a covariates	14.8	1	1	1	1	1.38	0.08	0.63	1.16	1.39	1.63	1.99
Estimate: sub-sample of males	1.4	1	1	1	0	1.74	0.12	0.94	1.39	1.75	2.10	2.60
Estimate: sub-sample of youths	1.3	1	1	1	1	−1.03	0.04	−1.31	−1.12	−1.03	−0.94	−0.37
Estimate: sub-sample of urban area	0.2	1	1	1	0	1.40	0.13	0.37	0.99	1.40	1.81	2.73
Estimate: sub-sample with high detection probability	0.3	1	1	1	1	−2.10	0.11	−3.01	−2.48	−2.10	−1.73	−1.34
Estimate: exogenous, index multiplicative	1.9	1	1	1	1	1.09	0.06	0.20	0.93	1.08	1.28	1.52
Estimate: exogenous, index miscellaneous	0.8	1	1	0	0	1.41	0.16	0.03	0.92	1.42	2.09	2.96
Estimate: exogenous, index items unprocessed	2.8	1	1	1	1	1.20	0.07	0.38	1.00	1.19	1.45	1.85
Estimate: study type, death penalty	8.2	1	1	1	1	3.23	0.15	2.16	2.77	3.25	3.60	5.31
Estimate: exogenous, crime data, arrest rate	9.5	1	1	1	1	−3.02	0.13	−3.94	−3.43	−3.00	−2.68	−2.00
Estimate: exogenous, crime data, conviction rate	3.9	1	1	1	1	−1.56	0.06	−1.90	−1.72	−1.55	−1.42	−1.10
Estimate: exogenous, crime data, incarceration rate	0.6	1	1	0	0	0.67	0.08	0.08	0.46	0.68	0.93	1.37
Estimate: exogenous, crime data, mean sentence length (sentenced)	2.0	1	1	1	0	1.83	0.15	0.91	1.48	1.84	2.59	3.27
Estimate: exogenous, crime data, mean sentence length (served)	1.7	1	1	0	0	1.65	0.18	0.04	1.08	1.67	2.13	2.87
Estimate: exogenous, crime data, police expenditures	4.0	1	1	1	1	1.82	0.08	1.30	1.61	1.82	2.10	2.76
Estimate: exogenous, crime data, police strength	7.9	1	1	1	1	3.09	0.08	2.62	2.88	3.09	3.33	3.89
Estimate: exogenous, crime data, other	8.5	1	1	1	0	0.62	0.07	0.20	0.43	0.63	0.80	1.50
Estimate: exogenous, crime data, probability dummy (regime shift)	3.5	1	1	1	0	−2.72	0.23	−5.05	−3.28	−2.73	−2.01	−0.70
Estimate: exogenous, crime data, severity dummy (regime shift)	3.6	1	1	1	1	−1.55	0.08	−2.28	−1.78	−1.55	−1.30	−1.00
Estimate: exogenous, crime data, incarceration per crime	0.9	1	1	1	1	−2.06	0.12	−3.32	−2.46	−2.05	−1.62	−1.08
Estimate: exogenous, crime data, convicted per crime	0.6	1	1	1	1	−2.40	0.10	−3.51	−2.69	−2.39	−2.16	−1.39
Estimate: exogenous, survey, type of punishment	0.3	1	1	1	0	1.41	0.11	0.58	1.13	1.40	1.79	2.39
Estimate: exogenous, survey, probability of other kind of punishment	0.4	1	1	0	0	2.32	0.24	0.34	1.50	2.33	3.06	4.26
Estimate: exogenous, survey, probability of detection by friends or family	0.2	1	1	1	1	−2.74	0.18	−4.07	−3.23	−2.76	−2.12	−1.45
Estimate: exogenous, survey, probability of punishment by friends or family	1.6	1	1	1	1	−3.44	0.20	−4.71	−4.00	−3.46	−2.78	−2.05
Estimate: exogenous, survey, severity of punishment by friends or family	1.0	1	1	0	0	1.53	0.19	0.44	0.99	1.52	2.13	2.96
Estimate: exogenous, survey, probability of detection by others	0.4	1	1	1	1	1.43	0.05	0.92	1.26	1.42	1.63	1.88
Estimate: exogenous, survey, probability of punishment by others	0.9	1	1	1	0	1.37	0.15	0.52	0.95	1.35	1.87	2.67
Estimate: exogenous, survey, severity of punishment by others	0.4	1	1	1	1	1.30	0.04	0.86	1.18	1.29	1.44	1.62
Estimate: exogenous, survey, time between offense and clearance	0.1	1	1	1	0	−1.09	0.09	−1.86	−1.43	−1.08	−0.87	−0.61
Estimate: exogenous, in differences	3.1	1	1	1	1	2.33	0.13	1.36	2.03	2.34	2.80	4.14
Estimate: endogenous, number of convicted	0.3	1	1	0	0	1.61	0.17	0.20	1.09	1.63	2.09	5.04
Estimate: endogenous, probability of delinquency of fictitious offense (surveyed is delinquent)	1.2	1	1	0	0	1.60	0.20	0.26	1.01	1.59	2.15	3.09
Estimate: endogenous and exogenous relate not to the same offense	6.3	1	1	1	1	1.96	0.12	1.34	1.66	1.96	2.35	4.20
Estimate: crime category, other	1.5	1	1	1	0	2.13	0.17	0.55	1.72	2.12	2.84	3.73
Estimate: offense, manslaughter	1.2	1	1	0	0	0.72	0.10	0.05	0.42	0.72	1.14	1.87
Estimate: offense, negligent assault	1.8	1	1	0	0	0.80	0.10	0.25	0.55	0.80	1.14	1.86
Estimate: offense, burglary	12.2	1	1	1	0	1.36	0.09	0.62	1.05	1.36	1.62	1.95

continued on the next page. . .

…last page of table 3.51 continued

Variable	var	A	B	C	D	mean	sd	min	1%	50%	99%	max
Estimate: offense, drug possession (soft)	0.7	1	1	1	0	1.38	0.11	0.59	1.12	1.37	1.82	2.51
Estimate: offense, drug possession (hard)	0.5	1	1	1	1	−1.26	0.03	−1.85	−1.40	−1.26	−1.21	−0.98
Estimate: offense, sexual assault	0.5	1	1	1	0	1.40	0.10	0.53	1.11	1.41	1.72	2.44
Estimate: offense, embezzlement	0.1	1	1	1	1	−1.62	0.09	−3.01	−1.96	−1.62	−1.37	−1.02
Estimate: offense, other crimes	8.1	1	1	0	0	1.10	0.10	0.09	0.85	1.11	1.41	1.80
Estimate: offense, crime rate (general)	6.2	1	1	1	0	1.15	0.09	0.54	0.88	1.17	1.36	1.67
Estimate: offense, vehicle theft	8.5	1	1	1	0	1.32	0.09	0.59	1.04	1.32	1.56	1.86
Estimate: offense, environmental crimes, Violations of prescriptive limits	2.3	1	1	1	1	1.47	0.07	0.72	1.25	1.47	1.71	2.10
Estimate: violent characteristics	15.1	1	1	1	0	−1.40	0.10	−2.36	−1.66	−1.40	−1.13	−0.38
Estimate: endogenous, metric category	20.0	1	1	0	0	2.27	0.19	0.07	1.68	2.28	2.76	3.29
Estimate: endogenous, interval category	4.1	1	1	0	0	1.19	0.17	0.25	0.75	1.18	1.75	2.85
Estimate: endogenous, binary category	9.6	1	1	1	1	−2.15	0.11	−2.71	−2.40	−2.16	−1.82	−0.62
Estimate: endogenous, other transformation	7.8	1	1	1	1	−1.20	0.06	−1.76	−1.42	−1.20	−1.08	−0.38
Estimate: endogenous and exogenous relate to the same time	28.0	1	1	0	0	−0.98	0.12	−1.97	−1.29	−0.99	−0.62	−0.05
Estimate: endogenous occurs before exogenous (lagged endogenous)	1.3	1	1	1	1	2.05	0.07	1.25	1.86	2.05	2.27	2.55
Estimate: covariate, sex	14.3	1	1	0	0	−0.57	0.09	−1.29	−0.87	−0.57	−0.32	−0.04
Estimate: covariate, nationality	3.3	1	1	1	1	2.73	0.12	1.94	2.41	2.74	3.09	3.80
Estimate: covariate, profession	0.4	1	1	0	0	0.83	0.09	0.11	0.51	0.84	1.06	1.77
Estimate: covariate, social integration	1.5	1	1	1	0	2.05	0.17	0.64	1.52	2.04	2.59	3.28
Estimate: covariate, religion	1.7	1	1	1	0	1.31	0.12	0.39	0.96	1.31	1.66	2.23
Estimate: covariate, drug usage	1.3	1	1	1	1	−0.88	0.06	−1.30	−1.12	−0.88	−0.75	−0.59
Estimate: covariate, previous convictions	1.6	1	1	1	1	2.93	0.18	1.65	2.32	2.93	3.50	4.18
Estimate: covariate, personal characteristics	3.0	1	1	1	1	−1.51	0.05	−1.90	−1.66	−1.51	−1.33	−1.01
Estimate: covariate, fixed effects (spatial)	10.0	1	1	1	1	−1.74	0.09	−2.71	−2.12	−1.73	−1.54	−1.29
Estimate: covariate, random effects	1.0	1	1	1	1	2.48	0.11	1.30	2.07	2.50	2.77	3.29
Estimate: covariate, other	35.8	1	1	1	0	2.08	0.17	0.88	1.53	2.11	2.58	3.09
Estimate: covariate, time trend	5.4	1	1	0	0	1.60	0.15	0.30	1.16	1.62	1.97	2.77
Estimate: covariate, poverty, welfare	6.4	1	1	1	1	3.25	0.10	2.37	2.94	3.26	3.51	3.95
Estimate: covariate, urbanity	8.2	1	1	1	0	1.92	0.15	1.15	1.60	1.92	2.57	3.99
Estimate: covariate, GDP	1.4	1	1	1	0	−1.04	0.07	−1.72	−1.24	−1.03	−0.85	−0.47
Estimate: covariate, population (-growth)	11.5	1	1	1	0	1.43	0.11	0.57	1.10	1.44	1.76	2.21
Estimate: covariate, alcohol (consumption)	1.7	1	1	1	1	2.41	0.12	1.47	2.06	2.42	2.69	3.39
Estimate: covariate, labor force	2.5	1	1	1	1	−1.65	0.10	−2.41	−1.94	−1.64	−1.38	−0.77
Estimate: covariate, risk propensity	0.8	1	1	1	0	4.67	0.33	2.51	3.53	4.68	5.85	7.37
Estimate: no correction for simultaneity	19.3	1	1	1	1	−2.09	0.12	−3.08	−2.44	−2.11	−1.76	−0.58
Estimate: correction for simultaneity (with methodology)	9.1	1	1	1	1	2.13	0.12	0.68	1.87	2.14	2.49	3.20
Estimate: weighted model	6.8	1	1	1	1	1.13	0.06	0.56	0.98	1.12	1.32	1.69
Estimate: unweighted model	6.8	1	1	1	1	−1.13	0.06	−1.69	−1.32	−1.12	−0.98	−0.56
Estimate: bivariate method, bivariate regression	0.6	1	1	1	1	−2.70	0.12	−3.26	−2.96	−2.71	−2.25	−1.72
Estimate: bivariate method, correlation	9.5	1	1	1	1	2.43	0.16	1.21	2.00	2.43	2.84	3.82
Estimate: bivariate method, point biserial correlation	0.4	1	1	1	0	−1.33	0.12	−2.80	−1.88	−1.34	−1.04	−0.64
Estimate: bivariate method, ρ	0.5	1	1	1	0	−1.38	0.15	−3.01	−2.05	−1.37	−1.03	−0.43
Estimate: bivariate method, t-test for independent samples	0.5	1	1	1	1	−1.57	0.07	−3.16	−1.77	−1.56	−1.40	−1.24
Estimate: bivariate method, τ	0.9	1	1	1	1	−2.42	0.13	−3.33	−2.79	−2.42	−2.08	−1.29

continued on the next page…

. . . last page of table 3.51 continued

Variable	var	A	B	C	D	mean	sd	min	1%	50%	99%	max
Estimate: multivariate method, OLS	24.2	1	1	1	0	1.83	0.15	0.48	1.35	1.85	2.27	3.01
Estimate: multivariate method, GMM	0.8	1	1	1	1	3.06	0.20	0.79	2.40	3.09	3.61	4.42
Estimate: multivariate method, other ML method	1.1	1	1	1	0	2.77	0.20	0.73	2.11	2.80	3.49	4.34
Estimate: multivariate method, GLS	1.7	1	1	1	1	−3.20	0.14	−3.91	−3.55	−3.21	−2.62	−2.07
Estimate: multivariate method, VAR	0.6	1	1	1	0	−1.05	0.10	−2.19	−1.35	−1.04	−0.74	−0.44
Estimate: test of significance	5.5	1	1	1	0	0.95	0.08	0.37	0.78	0.95	1.15	2.17
Estimate: no test of significance	5.5	1	1	1	0	−0.95	0.08	−2.17	−1.15	−0.95	−0.78	−0.37
Estimate: square root of sample size for negative values	79.2	1	1	1	1	−3.38	0.07	−4.35	−3.67	−3.37	−3.26	−3.04
Estimate: square root of sample size for positive values	82.5	1	1	1	1	6.21	0.07	5.76	6.06	6.20	6.44	6.98

A, B, C and D correspond to the four tests of significance described in subsection 3.6.4. The column *var* refers to the variation of a variable (i.e., the percentage of valid observations); the maximum variation for dummy variables is fifty percent. All other values are properties of the distribution of the t-values in the regressions. For reasons of parsimony we report only those variables in this table, which pass more than one test; 244 variables which only pass test A or are not considered stable are not shown (in fact, only 114 out of 515 variables do not pass test A). The reference category for dummies is usually the opposite property or, in the case of multiple categories, the missing values.

end of the table 3.51

The results concerning the countries of the authors are, at large, compatible with the bivariate analysis in section 3.5. Canadian, Australian, Swedish and authors from Switzerland have more positive t-values, while authors from the U.S., UK and Finland[48] have more negative t-values. The results for the author-variables, which pass at least one of the tests, are also in line with table 3.36 with the exception of Nagin, who has now a positive effect on the (normalized) t-values and Marvell who has a negative effect.

Surprisingly, the only disciplines of the authors to pass the tests are psychology and miscellaneous; the direction of their influences are as expected (table 3.30 seemed to suggest that being an economist, sociologist or criminologist would also be important). Nevertheless, these effects are found in the characteristics of the publisher and the institutional background of the authors. Authors from an economic or law institution present more negative results, while authors with a criminological or miscellaneous background seem to have the opposite effect. The same applies to the type of the publisher.

Looking at the specific journals delivers a more diversified picture: while all significant criminological journals have rather positive values, economical, sociological and other journals can be found on both sides. The number of selected journal shows no obvious tendency towards a specific discipline.

It is also noteworthy that journals are associated with rather positive (normalized) t-values, while books and working papers are associated with a negative effect. Assuming that a publication bias is, if at all, more present for studies published in journals, this would imply that the results are

[48] In fact, only Matti Virén comes from Finland.

slightly biased towards zero. However, recalling table 3.24 casts some doubts on this reasoning because analytical evidence indicates a publication bias for books and working papers but not for journal articles.

Those studies which examine Canadian data, use cross sections, employ a miscellaneous public data base, present large numbers of bivariate estimates or study the states of a nation, stand out and have a positive effect on the (normalized) t-values. The opposite effect is found among studies which employ complete samples, experiments, data from the United Kingdom, Australia or Germany, use data from large cities, report checks of significance or checks for plausibility and errors.

Looking at the implemented covariates, the results of table 3.45 are largely replicated. Controlling for spatial fixed effects, labor force, GDP, drug usage and personal characteristics imply more negative (normalized) t-values. Only the latter does not comply with the results from the ANOVA. Smaller (i.e., more positive) effects are found when religion, population, urbanity, social integration, alcohol consumption, random effects, nationality, previous convictions, poverty and time trends are controlled for. Again, only the latter does not conform with the ANOVA.

There are only a few studied crimes which have a significant influence on the outcome of an estimate. Of those with a negative influence, no one passes more than one test. Malicious mischief, speeding, severe larceny and drunk driving are compatible with the results from table 3.43 but pass only the first test, while general drug related crimes have a negative effect (opposite to the results from the ANOVA). Among the crimes with a positive effect, all are in line with the results from the ANOVA and the following crimes pass more than one test: burglary, general crime rate, environmental offenses, manslaughter, negligent assault, other crimes and vehicle theft. Environmental offenses stand out by passing all four tests while the effect is ambiguous in the ANOVA.

The results of the variables which measure deterrence comply with table 3.42. Especially the arrest and conviction rate and the regime shift dummies have a negative effect and pass three or four tests. A positive effect is found among the sentence lengths and police related deterrence variables. Among the implemented methods, the usage of correlations, OLS and 2SLS have a very prominent positive effect on the (normalized) t-values, while ARIMA- and GLS-methods imply more negative values. These are more or less in line with table 3.48.

It is also interesting to note that exogenous variables in binary form are negatively associated, while metric variables and those measuring intervals have a positive (normalized) impact on t-values. More negative (normalized) t-values are found when simultaneity is not accounted for, the model is not weighted, deterrence is the focus of a study, a sample of youths is studied, the exogenous variables are transformed in another form (than log or differences) or a violent crime is studied. A positive effect is found when the deterrent variable and the endogenous variable do not relate to the same offense, the number of included covariates increases, simultaneity is methodologically accounted for, the deterrence variable is measured in differences or the study is about the death penalty. The most significant variable is the square root of the sample size, diversified by the sign of the t-values. This relationship indicates the existence of an effect and is described in section 3.4.

3.6.5 Stepwise Regressions

If I should throw down a thousand beans at random upon a table, I could doubtless, by eliminating a sufficient number of them, leave the rest in almost any geometrical pattern you might propose to me, and you might then say that that pattern was the thing prefigured beforehand, and that the other beans were mere irrelevance and packing material. Our dealings with Nature are just like this.

William James, The Varieties of Religious Experience, 1902

The idea of stepwise regressions is to include (or exclude) variables in a regression as long as they enhance the regression-model. The degree of improvement can be measured by different means. Some algorithms use the R^2, the Bayesian or Akaike Information Criterion (BIC or AIC), or other properties to measure the overall model-quality. However, other algorithms do not rely on overall model properties but resort to attributes of the individual variables (e.g., the significance of a variable).

We use the sw-algorithm implemented in STATA and the stepAIC-algorithm in R (R Development Core Team, 2007). The first is based on the significance of the variables while the latter is based on the improvement of the model-fit (measured by the AIC). Both algorithms are used in conjunction with OLS and with the forward and backward strategy. Besides the usage in many other fields, the principle of stepwise regressions has been applied in empirical deterrence studies like Allison (1972); Cho (1972); Cloninger (1975); Gross and Hakim (1982); Norström (1983); Meera and Jayakumar (1995); Velez et al. (1999); Freeman et al. (2006) and many more.

The sw-Algorithm

The sw-algorithm in- and excludes variables according to their significance in a model. In each step, variables are either included if they are highly significant or excluded if they are not significant. Therefore, the choice of these two thresholds p_1 and p_2 are decisive. Lovell (1983) shows that in the case of selecting the "best" variables from a large set of possible candidates, it does not suffice to resort to the standard values of significance. When the pool of variables increases, the set of those variables which are significant by chance will increase as well. We have more than 500 variables. Even if these are all independent and random, we would expect at least 25 variables to pass the five percent significance threshold. However, in the search of the most influencing variables we can circumvent this problem by reducing the threshold-values accordingly.

Lovell (1983) gives a handy formula to adjust those threshold values: $\hat{\alpha} := p \cdot \frac{c}{k}$, where c is the number of variables at hand and k is the number of desired significant variables. We have 515 variables out of which 493 can be used simultaneously in a OLS-regression (some have to be dropped due to singularity problems). A first regression, using all variables determined 44 variables to be significant at a five percent level. Since k must be arbitrarily chosen, we take $k = 50$ for the inclusion- and a less conservative choice of $k = 80$ for the exclusion threshold and set $c = 493$. So, instead of using 0.05 and 0.2 as the in- and exclusion probabilities, we try to lessen the selection bias by using the adjusted thresholds of $p_1 = 0.05\frac{50}{493} \approx 0.005$ and $p_2 = 0.2\frac{80}{493} \approx 0.03$.

Hendrey and Krolzig (2000) show that these adjustment can be avoided when the data mining algorithm accounts for the selection bias as is the case with their software PcGETS. However, this accumulation of statistical tools didn't achieve any better results for a restricted test set of 200 studies[49]. Therefore, we resort to the well studied stepwise regression methods.

Backward Stepwise Regression

The basic procedure of the stepwise regression algorithms implemented in STATA is rather simple. Both methods, backward and forward are very similar.

1. start with the full model,
2. exclude the least significant variable if its p-value is above p_2,
3. include the most significant excluded variable if its p-value is below p_1,
4. exclude the least significant included variable if its p-value is above p_2,
5. re-estimate and repeat steps 3 to 4 until neither is possible.

Forward Stepwise Regression

The forward procedure typically finds less variables than the backward procedure but is methodologically almost identical:

1. start with the empty model,
2. include the most significant variable if its p-value is below p_1,
3. exclude the least significant included variable if its p-value is above p_2,
4. include the most significant excluded variable if its p-value is below p_1,
5. repeat steps 3 to 4 until neither is possible.

The stepAIC-Algorithm

The applied algorithm is similar to the sw-algorithm. It adds and drops variables but evaluates the changes in the model-fit (the Aikaike Information Criterion) instead of the significance values. The algorithm stops as soon as any further improvement is smaller than a certain threshold-value; see also Venables and Ripley (2002) for more detailed information.

We use the implementation in R with the default settings. Each algorithm is applied to the whole data set (backward) and to the empty data set (forward). The results of these two algorithms are given in table 3.52.

[49] One reason may be that it does fit time series models better than our data set of very heterogenous cross sections.

Table 3.52: Multivariate analysis - stepwise regressions

Variable	var	1c	1t	2c	2t	3c	3t	4c	4t
		sw				stepAIC			
Study: publication, page begin	99.8					0.01	4.0	0.01	3.9
Study: publication, page end	100.0					−0.01	−4.3	−0.01	−4.1
Study: not explorative	4.0					−0.49	−2.6	−0.44	−2.4
Study: measuring points	97.4					0.00	3.1	0.00	3.1
Study: year of first measure	84.4							0.00	−2.9
Study: time span in months	63.2					−0.00	−3.0	−0.00	−2.9
Study: size of first population	24.9	0.00	2.9			0.00	2.3		
Study: size of second population	0.9	0.00	4.8			0.00	3.8	0.00	1.6
Study: size of first sample	25.6					0.00	−3.7	0.00	−1.7
Study: size of second sample	1.2	0.00	4.9					0.00	−1.7
Study: size of first realized sample	58.8	0.00	−3.8	0.00	−4.9	0.00	−9.2	0.00	−7.8
Study: size of second realized sample	2.4	−0.00	−14.1	−0.00	−7.6	−0.00	−7.7	0.00	−5.5
Study: rate of return of first sample	13.0							0.00	1.6
Study: rate of return of second sample	0.9							−0.02	−2.3
Study: maximum age in first sample	5.7					−0.02	−4.3	−0.02	−3.5
Study: mean age in first sample	5.0					0.02	3.0	0.02	1.8
Study: check for validity	2.1							−0.43	−1.4
Study: tests of significance	9.4	−0.67	−3.4	−0.59	−2.6	−0.58	−4.5	−0.64	−4.8
Study: number of bivariate estimates	32.4			0.01	3.2	0.01	3.3	0.01	3.0
Study: user, tr	48.6			−1.11	−5.0	−0.53	−3.4	−0.65	−3.8
Study: user, aw	29.3					0.42	3.1	0.45	3.1
Study: user, mw	1.4					0.77	3.5	0.48	2.0
Study: publication, journal article	13.7					0.53	3.0	0.36	1.9
Study: publication, working paper, report	4.9					−0.97	−4.1	−1.11	−4.6
Study: publication, miscellaneous type	3.3					−0.48	−1.7	−0.60	−2.0
Study: publication, not a dissertation or master thesis, etc.	2.3					−1.63	−4.3	−1.53	−3.8
Study: author, Steven D. Levitt	1.7					1.94	6.3	1.75	5.3
Study: author, William C. Bailey	2.7					0.62	2.8		
Study: author, David W. Rasmussen	1.3					−1.45	−4.6	−1.85	−5.5
Study: author, Theodore G. Chiricos	1.2					0.64	1.7		
Study: author, Dale O. Cloninger	1.5					0.81	2.9		
Study: author, Simon Hakim	1.0					−1.36	−4.1	−1.74	−4.9
Study: author, Raymond Paternoster	2.4					−0.76	−2.5	−0.83	−2.7
Study: author, Isaac Ehrlich	1.0	−1.63	−4.1	−1.70	−2.8	−0.71	−2.0	−1.17	−3.2
Study: author, Matti Virén	0.8					−1.54	−3.4	−2.54	−5.8
Study: author, Ann Dryden Witte	1.0					0.71	1.8		
Study: author, Maynard I. Erickson	0.6			1.06	3.9			1.27	1.7
Study: author, Jack P. Gibbs	0.9							−1.42	−2.4
Study: author, Alex R. Piquero	1.3					−0.69	−2.0	−1.33	−3.6
Study: author, other author	22.5							−0.38	−3.2
Study: journal, Criminal Justice	2.2							0.54	2.0
Study: journal, Criminal Law and Criminology	2.2					−0.41	−1.8		

continued on the next page…

…last page of table 3.52 continued

			sw				stepAIC		
Variable	var	1c	1t	2c	2t	3c	3t	4c	4t
Study: journal, Social Forces	2.2					−0.47	−1.8	−0.77	−2.9
Study: journal, Law and Society Review	2.8							0.55	2.4
Study: journal, American Journal of Economics and Sociology	1.7							0.50	1.9
Study: journal, Public Economics	1.1							1.07	2.4
Study: journal, Social Problems	1.6							0.53	1.7
Study: journal, Accident Analysis and Prevention	2.2					−0.82	−3.1	−0.98	−3.2
Study: journal, Studies on Alcohol	1.1							−0.99	−2.6
Study: journal, Criminology	4.5					0.43	2.5	0.51	2.8
Study: journal, Social Science Quarterly	2.2			−0.84	−4.3	−0.60	−2.5	−0.47	−1.9
Study: journal, Legal Studies	2.0					−0.49	−2.0		
Study: publication, United Kingdom	8.3					0.60	4.3	0.69	4.8
Study: publication, Canada	3.6							−0.83	−2.8
Study: publication, Netherlands	2.6					0.61	2.8	0.64	2.9
Study: author, Germany	3.2					2.51	4.5	0.91	2.5
Study: author, USA	21.3							0.27	1.7
Study: author, Switzerland	1.0			1.41	4.6	3.25	5.2	3.35	5.3
Study: author, Finland	0.8	−0.43	−3.1	−0.97	−2.7				
Study: author, Netherlands	1.3			1.50	2.6	1.02	3.0	1.56	3.8
Study: author, Australia	2.0					0.68	2.6	1.02	2.9
Study: author, Sweden	0.9			2.02	4.3	2.73	7.1	2.79	5.4
Study: author, other country	3.7					0.73	3.7	0.82	3.4
Study: author, criminology	11.3					0.63	5.1	0.61	4.3
Study: author, psychology	4.2					−0.43	−2.1	−0.58	−2.4
Study: author, law	3.6					0.27	1.5		
Study: publication, economics	36.0					−0.58	−5.5	−0.46	−4.3
Study: publication, type not applicable	0.4	−1.40	−3.5	−2.07	−3.6	−2.67	−4.6	−2.97	−4.9
Study: publication, criminology	20.5	1.03	6.4						
Study: publication, sociology	19.4	0.70	3.9						
Study: publication, miscellaneous	14.1					−0.44	−3.3		
Study: publication, psychology	2.8					−1.42	−5.3	−0.82	−2.8
Study: institute, sociology	21.3							−0.19	−1.6
Study: institute, miscellaneous	14.9	0.88	4.2	0.74	3.7	0.94	8.5	0.86	7.2
Study: cross section	23.2			0.48	2.7	0.39	3.8	0.29	2.6
Study: single survey	17.8			−0.51	−2.8			−0.73	−3.1
Study: repeated survey	3.1					0.93	3.6	0.45	1.5
Study: panel survey	5.6							−0.87	−3.2
Study: experiment (laboratory)	4.4							−1.47	−4.4
Study: experiment (field, researcher initiative)	1.6					1.39	4.3		
Study: experiment (field, institutional initiative)	1.6					−0.97	−3.4	−1.23	−3.8
Study: experiment (natural)	3.0					0.47	2.3		
Study: not experimental	15.9							−0.60	−2.1
Study: quasi experimental	8.8							−0.63	−2.2
Study: first population, Germany	2.7					−1.98	−3.6		

continued on the next page…

...last page of table 3.52 continued

Variable	var	sw				stepAIC			
		1c	1t	2c	2t	3c	3t	4c	4t
Study: first population, United Kingdom	5.0					0.76	3.0		
Study: first population, USA	23.3			0.61	2.3	1.00	4.8		
Study: first population, Canada	4.3	0.56	2.7	1.39	4.6	2.04	8.3	1.98	7.9
Study: first population, Sweden	1.4							−1.00	−2.2
Study: first population, Finland	0.9	−0.42	−2.6						
Study: first population, Switzerland	1.0	0.61	2.8			−1.24	−1.8	−1.40	−2.2
Study: first population, Australia	1.7							−1.10	−3.0
Study: first population, Netherlands	1.3							−1.22	−2.8
Study: first population, other country	6.2			1.31	3.4	1.48	5.7	0.56	3.2
Study: second population, other country	3.3					0.97	2.4		
Study: sample base, first population, complete country	36.9					−1.82	−3.4	−1.77	−3.2
Study: sample base, first population, partial country	37.4					−2.16	−4.2	−2.05	−3.8
Study: sample base, second population, complete country	2.4			−3.60	−9.2	−2.99	−5.5	−2.22	−4.0
Study: sample base, second population, partial country	2.2			−3.51	−5.6	−3.95	−5.9	−3.78	−5.4
Study: sample unit, first population, states	21.9			0.45	2.6	0.40	4.2	0.37	3.8
Study: sample unit, first population, miscellaneous	7.4					0.97	6.5	0.92	6.0
Study: sample unit, second population, individuals	1.7					1.78	3.0	2.72	4.3
Study: sample individuals, first population, population	36.2					0.40	3.7	−0.51	−4.0
Study: sample individuals, first population, students	11.7							−0.55	−3.1
Study: sample individuals, first population, pupils	3.0					1.53	6.7	1.00	3.7
Study: sample individuals, second population, population	2.6			4.28	8.2	4.70	7.7	3.68	5.8
Study: sample individuals, second population, miscellaneous	1.2			4.71	6.5	3.70	5.4	3.36	4.4
Study: sample individuals, second population, students	0.5	−3.44	−3.5	1.80	2.9				
Study: sample individuals, second population, pupils	0.2			3.44	6.0	1.59	1.4		
Study: sample of extreme groups	0.5					1.21	2.2		
Study: complete sample	9.8					−0.64	−5.2	−0.73	−5.8
Study: PKS is public data base	1.2			1.07	2.6	3.45	7.0	2.35	4.6
Study: miscellaneous public data base	40.1					0.96	6.9	0.83	5.6
Study: UCR is public data base	22.9					0.63	4.0	0.56	3.4
Study: no public data base	26.3					0.61	4.3	0.66	4.5
Study: income representative	1.1					0.58	1.5		
Study: education below average	0.1							−1.69	−1.4
Study: no class overrepresented	0.9							1.62	3.2
Study: no social fringe group	0.6			−0.99	−3.1	−2.13	−3.6	−1.65	−2.7
Study: percentage of convicted $>75\%$	0.6			1.50	4.1	2.02	3.7	1.31	2.3
Study: percentage of convicted $>51-75\%$	0.3	−1.28	−2.9			1.13	1.8		
Study: main location >500000 inhabitants	3.5			−1.68	−2.3	−1.84	−9.1	−1.87	−8.9
Study: main location <5000 inhabitants	0.2	−3.04	−10.2	−2.80	−4.9	−7.38	−7.7	−6.92	−7.0
Study: small cities overrepresented	1.1					1.00	2.6	0.74	1.8
Study: does not claim to be representative	32.5					−0.36	−4.0	−0.26	−2.9
Study: claims to be representative	19.4					−0.57	−4.5	−0.47	−3.5
Study: does not check representativeness	26.9			0.47	3.0	0.45	4.8	0.49	5.0
Study: checks representativeness	2.9	−0.63	−3.1						

continued on the next page...

…last page of table 3.52 continued

			sw				stepAIC		
Variable	var	1c	1t	2c	2t	3c	3t	4c	4t
Study: does not report representativeness checks	2.2			−0.80	−3.2	−0.75	−2.5	−0.92	−3.0
Study: reports representativeness checks	0.4					−1.34	−2.4		
Study: closed questions for pretest	21.5			1.10	4.2	0.81	4.5	1.20	5.9
Study: mixed questions for pretest	2.1			2.06	5.0	1.43	4.5	1.56	4.7
Study: open questions for pretest	1.4					−1.04	−3.2	−0.76	−2.2
Study: Guttman reliability method	0.2					5.74	4.0	5.82	4.0
Study: miscellaneous reliability method	0.3			−4.42	−4.0	−3.87	−4.2	−4.45	−4.4
Study: correlational reliability method	0.2	−1.77	−10.6	−1.18	−5.5	−2.91	−3.6	−3.52	−4.2
Study: variables reliable	3.9					1.06	4.5	1.22	4.6
Study: validity test of all variables	0.2							3.88	2.9
Study: miscellaneous validity check	0.4							−1.85	−1.6
Study: criteria validity	1.0			1.32	2.7				
Study: the variables are not valid	0.2	−1.64	−3.8						
Study: conditions for significance check fulfilled	49.2							0.15	1.6
Study: conditions for significance check not fulfilled	3.6					0.52	2.6	0.44	1.9
Study: quality index	100.0					−0.09	−3.7	−0.06	−2.4
Estimate: exogenous, number of categories	35.3					0.00	2.0		
Estimate: endogenous, begin of observation (year)	78.5							0.00	2.2
Estimate: endogenous, number of categories	17.3							−0.01	−1.8
Estimate: deterrence is focus-variable	14.8					−0.30	−2.8	−0.35	−3.4
Estimate: complete sample	14.9					−0.20	−1.9	−0.17	−1.7
Estimate: sub-sample of males	1.4					−0.51	−1.7	−0.50	−1.6
Estimate: sub-sample of adults	0.5							−0.66	−1.4
Estimate: sub-sample of youths	1.3					−0.59	−1.8	−0.68	−2.0
Estimate: sub-sample of non-urban area	0.8	−1.02	−3.0			−1.79	−4.2	−1.68	−3.8
Estimate: exogenous, index multiplicative	1.9					1.03	3.1		
Estimate: exogenous, index additive	0.6							−1.05	−2.1
Estimate: exogenous, index mean	0.0	−1.15	−3.8	−3.76	−5.4				
Estimate: exogenous, index items unprocessed	2.8			1.19	3.1	0.43	1.5	1.21	5.5
Estimate: exogenous, index items miscellaneous	0.2							1.79	2.2
Estimate: exogenous, index items standardized	0.2	0.63	3.4					2.26	2.8
Estimate: study type, crime data	45.3					−1.13	−8.8	−1.46	−4.3
Estimate: study type, survey	24.4					−2.13	−2.9	−3.10	−3.6
Estimate: study type, experiment	12.7							−2.65	−2.8
Estimate: study type, death penalty	8.2			1.20	4.5				
Estimate: exogenous, death penalty, existence of death penalty	1.0	1.26	3.6			1.03	3.1	0.65	1.4
Estimate: exogenous, death penalty, execution rate	4.8							−0.52	−1.5
Estimate: exogenous, death penalty, other	1.5					−0.45	−1.6	−0.85	−2.0
Estimate: exogenous, crime data, clearance rate	5.3					0.56	3.3	0.57	3.2
Estimate: exogenous, crime data, arrest rate	9.5	−0.83	−3.4						
Estimate: exogenous, crime data, conviction rate	3.9					−0.36	−2.1	−0.39	−2.2
Estimate: exogenous, crime data, parole rate	0.2	−1.09	−4.4						
Estimate: exogenous, crime data, incarcerations (absolute or per capita)	0.9					1.61	4.7	1.50	4.3

continued on the next page…

…last page of table 3.52 continued

Variable	var	1c	1t	2c	2t	3c	3t	4c	4t
		sw				stepAIC			
Estimate: exogenous, crime data, incarceration rate	0.6					0.77	2.0	0.67	1.8
Estimate: exogenous, crime data, mean sentence length (sentenced)	2.0			0.89	3.9	1.06	4.6	1.15	5.0
Estimate: exogenous, crime data, mean sentence length (served)	1.7					0.73	2.9	0.77	3.0
Estimate: exogenous, crime data, police expenditures	4.0			1.43	3.3	1.54	8.3	1.48	7.8
Estimate: exogenous, crime data, police strength	7.9			1.53	4.1	1.44	10.1	1.37	9.6
Estimate: exogenous, crime data, other	8.5					0.69	5.0	0.63	4.5
Estimate: exogenous, crime data, probation rate	0.2					1.08	1.5	1.11	1.6
Estimate: exogenous, crime data, incarceration per crime	0.9					−0.61	−1.8	−0.57	−1.6
Estimate: exogenous, crime data, convicted per crime	0.6					−0.88	−2.2	−0.83	−2.1
Estimate: exogenous, survey, surveyed is delinquent	2.9							0.55	2.5
Estimate: exogenous, survey, is experiment	23.1					0.94	1.6	0.94	1.5
Estimate: exogenous, survey, is no experiment	0.9			2.43	2.7	3.01	4.3	2.66	3.6
Estimate: exogenous, survey, probability of detection by police	7.1			−0.85	−4.2	−0.65	−4.0	−0.75	−4.5
Estimate: exogenous, survey, probability of punishment by justice	4.5			−0.93	−4.6	−0.79	−4.0	−0.83	−4.3
Estimate: exogenous, survey, prob. of punishment by employment law	0.6			−0.87	−2.9				
Estimate: exogenous, survey, prob. of detection by friends or family	0.2	−1.10	−3.1	−2.24	−3.9	−1.51	−2.1	−1.41	−1.9
Estimate: exogenous, survey, prob. of punishment by friends or family	1.6	−0.87	−3.5	−1.52	−5.8	−1.21	−4.4	−1.34	−4.7
Estimate: exogenous, survey, probability of punishment by others	0.9					0.54	1.5		
Estimate: exogenous, survey, severity of punishment by others	0.4					1.49	3.0	1.35	2.7
Estimate: exogenous, survey, time between offense and clearance	0.1			−1.49	−2.8	−1.11	−1.4	−1.10	−1.4
Estimate: exogenous, survey, relates to the present	21.3			0.80	2.6	1.44	2.1	2.01	2.5
Estimate: exogenous, survey, relates to the past	2.7	1.08	4.0	1.13	3.5	1.92	2.6	2.60	3.2
Estimate: exogenous, experiment, yes	7.2							3.02	3.4
Estimate: exogenous, experiment, no	5.4			1.01	3.1			2.93	3.3
Estimate: exogenous, experiment, experimental variation of probability of detection	2.1			−1.72	−3.9	−1.42	−5.4	−1.17	−4.0
Estimate: exogenous, experiment, other	2.2					0.46	1.8	0.44	1.5
Estimate: exogenous, experiment, relates to the present	13.9							−0.79	−3.3
Estimate: exogenous, experiment, relates to the past	0.5							−1.37	−2.2
Estimate: exogenous, relates to one year	42.7			0.39	2.3	0.86	7.4	0.64	4.8
Estimate: exogenous, relates to more than one year	12.1							−0.47	−2.8
Estimate: exogenous, metric category	36.8							1.55	3.4
Estimate: exogenous, interval category	9.1							1.57	3.2
Estimate: exogenous, binary category	18.8							1.56	3.4
Estimate: exogenous, nominal category	0.3					−3.59	−3.1	−2.62	−2.1
Estimate: exogenous, ordinal category	7.4					0.52	3.0	1.90	3.9
Estimate: exogenous, in logs	20.2					−0.21	−1.5	−0.28	−2.0
Estimate: exogenous, in differences	3.1	0.77	3.2	2.01	3.4	1.73	7.0	1.73	7.1
Estimate: exogenous, not in differences	0.2					1.25	1.8		
Estimate: exogenous, not other transformation	10.1					−0.59	−3.9	−0.70	−4.3
Estimate: endogenous, index miscellaneous	0.1	3.06	6.2			2.09	1.9	2.06	1.9
Estimate: endogenous, index multiplicative	0.2			3.79	2.7	1.38	1.5	1.98	2.0
Estimate: endogenous, number of reported crimes (absolute numbers)	11.1							0.30	2.3

continued on the next page…

…last page of table 3.52 continued

		sw				stepAIC			
Variable	var	1c	1t	2c	2t	3c	3t	4c	4t
Estimate: endogenous, number of registered suspects	0.6					1.35	3.1	1.50	3.3
Estimate: endogenous, number of convicted to prison sentence	0.2			1.33	3.8	2.19	3.1	2.16	2.9
Estimate: endogenous, probability of future delinquency (surveyed is delinquent)	4.1					−0.31	−1.5		
Estimate: endogenous, probability of delinquency of fictitious offense (surveyed is delinquent)	1.2					0.68	2.0	1.36	3.6
Estimate: endogenous, recidivism	0.7			1.26	2.5	1.94	3.6	2.24	3.9
Estimate: endogenous, accidents	4.2			0.83	2.9	1.09	5.2	0.82	3.5
Estimate: endogenous, violating prescriptive limits	2.9					−0.71	−2.5		
Estimate: endogenous, relates to less than one year	29.2					0.58	4.7	0.45	3.3
Estimate: endogenous, relates to more than one year	11.5					0.22	1.6	0.51	3.0
Estimate: endogenous, lifelong self reported delinquency	3.8					−0.46	−2.3		
Estimate: endogenous, one year of self reported delinquency	8.2							0.41	2.3
Estimate: endogenous, self reported delinquency since age of fourteen	0.6					2.34	4.8	2.01	4.0
Estimate: endogenous, less than one year of unlimited future self reported delinquency	1.2							−0.56	−1.5
Estimate: endogenous and exogenous relate not to the same offense	6.3					0.58	3.8	0.69	4.4
Estimate: crime category, misdemeanors	9.5					−0.98	−6.6	−0.96	−5.9
Estimate: crime category, formal deviant behavior	2.5					0.75	3.1	0.77	3.0
Estimate: crime category, violation of game-rules	2.4							0.74	2.7
Estimate: crime category, other	1.5					0.73	2.4	0.95	2.9
Estimate: offense, assault	10.1					−0.24	−2.1	−0.29	−2.4
Estimate: offense, negligent assault	1.8					0.64	2.4	0.79	3.0
Estimate: offense, burglary	12.2					0.22	1.9	0.29	2.5
Estimate: offense, larceny (severe)	3.2					−0.37	−1.8	−0.57	−2.7
Estimate: offense, drug possession (soft)	0.7			2.58	3.0	2.27	5.2	2.42	5.0
Estimate: offense, drug possession (hard)	0.5			−3.27	−2.4	−3.01	−5.0	−3.39	−5.4
Estimate: offense, drug related crime (general)	4.4							−0.52	−2.6
Estimate: offense, other sexual related crimes	0.8					0.59	1.5		
Estimate: offense, speeding	1.1					0.74	2.2	0.72	1.8
Estimate: offense, drunk driving	12.1					0.30	2.0	0.46	2.6
Estimate: offense, fare dodging	0.4					1.19	2.0	1.03	1.8
Estimate: offense, fraud	3.9					0.34	1.8	0.31	1.5
Estimate: offense, tax evasion	7.3					0.49	2.8	0.53	2.9
Estimate: offense, embezzlement	0.1			−1.87	−2.9				
Estimate: offense, other	7.1					−0.70	−4.3	−0.83	−4.6
Estimate: offense, vehicle theft	8.5					0.35	2.6	0.28	2.1
Estimate: offense, environmental crimes, Violations of prescriptive limits	2.3					1.43	4.7	1.14	4.3
Estimate: property and violent characteristics	48.8			0.37	2.3	0.37	4.6	0.44	4.6
Estimate: violent characteristics	15.1							0.21	1.7
Estimate: endogenous, metric category	19.5							−1.16	−2.1
Estimate: endogenous, interval category	4.1							−0.85	−1.4
Estimate: endogenous, ordinal category	4.6					−1.25	−5.6	−2.46	−4.2

continued on the next page…

. . . last page of table 3.52 continued

Variable	var	sw				stepAIC			
		1c	1t	2c	2t	3c	3t	4c	4t
Estimate: endogenous, binary category	9.6			−0.63	−2.5	−1.17	−8.8	−2.46	−4.4
Estimate: endogenous, nominal category	0.3							−1.40	−1.6
Estimate: endogenous, not in logs	32.0					0.59	3.0	0.68	3.3
Estimate: endogenous, in logs	26.8					0.57	2.5	0.68	2.9
Estimate: endogenous, other transformation	7.8			−1.26	−2.4	−1.49	−9.3	−1.62	−9.7
Estimate: endogenous and exogenous relate to the same time	28.0					−0.57	−2.4	−0.92	−3.8
Estimate: endogenous occurs after exogenous (lagged exogenous)	25.6					−0.43	−1.7	−0.68	−2.7
Estimate: covariate, age	20.2					−0.33	−3.3	−0.35	−3.4
Estimate: covariate, marital status	5.6					0.43	2.5	0.53	3.1
Estimate: covariate, profession	0.4					0.87	1.6	1.04	1.9
Estimate: covariate, social integration	1.5					0.51	1.7	0.84	2.5
Estimate: covariate, religion	1.7					0.49	1.7	0.56	1.9
Estimate: covariate, social class	0.8					−1.06	−2.9	−1.51	−4.0
Estimate: covariate, drug usage	1.3					−1.22	−3.6	−1.38	−4.1
Estimate: covariate, acceptance of norms	2.2							−0.47	−1.7
Estimate: covariate, morality	1.6	0.52	2.9			1.05	3.5	1.00	3.1
Estimate: covariate, personal characteristics	3.0					−0.57	−2.5	−0.76	−3.1
Estimate: covariate, fixed effects (spatial)	10.0					−0.58	−4.4	−0.39	−2.8
Estimate: covariate, fixed effects (time)	12.0							−0.30	−2.4
Estimate: covariate, random effects	1.0					1.02	3.1	1.00	3.0
Estimate: covariate, other	35.8					0.15	1.6	0.21	1.8
Estimate: covariate, time trend	5.4					0.35	2.2	0.35	2.1
Estimate: covariate, poverty, welfare	6.4	1.03	3.5	0.97	3.3	0.69	4.8	0.73	4.9
Estimate: covariate, urbanity	8.2	0.80	3.2			0.41	3.2	0.44	3.4
Estimate: covariate, GDP	1.4					−0.98	−3.2	−1.04	−3.4
Estimate: covariate, population (-growth)	11.5					0.47	4.0	0.44	3.7
Estimate: covariate, alcohol (consumption)	1.7					0.67	2.5	0.92	3.2
Estimate: covariate, labor force	2.5					−0.34	−1.6	−0.35	−1.6
Estimate: covariate, consumption	2.0					−0.68	−2.7	−0.81	−3.2
Estimate: covariate, risk propensity	0.8	0.85	3.4	1.27	3.9	2.39	5.6	2.22	5.1
Estimate: not linear model	15.1							−0.41	−2.1
Estimate: additive model	6.3					0.26	1.4	0.35	1.5
Estimate: not additive model	1.6							0.81	2.1
Estimate: correction for simultaneity (with variables)	9.1							−0.40	−2.1
Estimate: no correction for simultaneity	19.3			−0.69	−3.3	−1.00	−8.9	−1.18	−7.6
Estimate: no error correction	12.8					0.22	2.0		
Estimate: weighted model	6.8					0.41	2.8	0.35	2.4
Estimate: bivariate method, bivariate regression	0.6			−1.36	−2.9	−1.30	−3.4	−1.80	−4.3
Estimate: bivariate method, other nonparametric test	0.1							−2.08	−1.8
Estimate: bivariate method, correlation	9.5	0.59	4.0					−0.52	−2.6
Estimate: bivariate method, differences (means, percentages, etc.)	4.0							−0.56	−2.3
Estimate: bivariate method, point biserial correlation	0.4			−0.69	−3.0	−1.02	−1.9	−1.59	−2.8
Estimate: bivariate method, ρ	0.5					−2.23	−4.3	−3.03	−4.6

continued on the next page. . .

... last page of table 3.52 continued

Variable	var	sw				stepAIC			
		1c	1t	2c	2t	3c	3t	4c	4t
Estimate: bivariate method, t-test for independent samples	0.5					−1.43	−2.7	−1.77	−3.2
Estimate: bivariate method, t-test for dependent samples	0.6							−0.99	−1.9
Estimate: bivariate method, τ	0.9					−0.64	−1.7	−1.33	−3.2
Estimate: bivariate method, ANOVA	2.1					0.82	3.2		
Estimate: bivariate method, other	1.0					1.13	3.2		
Estimate: bivariate method, Wilcoxon	0.2	−0.74	−2.8						
Estimate: bivariate method, binomial	0.2	1.51	10.2	0.85	3.6				
Estimate: multivariate method, OLS	24.2							−0.61	−3.2
Estimate: multivariate method, 2SLS	9.8			−0.60	−2.5	−0.64	−5.0	−1.39	−6.1
Estimate: multivariate method, GMM	0.8					0.78	2.0		
Estimate: multivariate method, poisson regression	1.1					−0.43	−1.4	−1.07	−2.9
Estimate: multivariate method, other ML method	1.1					1.11	2.9		
Estimate: multivariate method, other	5.1			−0.75	−2.3	−0.61	−3.9	−1.26	−6.6
Estimate: multivariate method, ANOVA	0.4	−1.69	−3.4	−1.57	−4.6			−1.69	−2.8
Estimate: multivariate method, GLS	1.7			−0.97	−2.5	−1.70	−6.2	−2.36	−7.0
Estimate: multivariate method, VAR	0.6							−1.07	−2.3
Estimate: multivariate method, path analysis	1.3					−0.60	−2.1	−1.30	−3.8
Estimate: multivariate method, ARIMA	4.2					−0.62	−3.4	−1.24	−5.0
Estimate: multivariate method, COX regression	0.3	1.89	8.2	2.63	4.9	1.82	2.7	1.66	2.4
Estimate: test of significance	5.5					0.54	3.0	0.52	2.8
Estimate: square root of sample size for negative values	79.2	−0.02	−4.2	−0.01	−4.5	−0.01	−12.5	−0.01	−13.3
Estimate: square root of sample size for positive values	82.5	0.05	6.9	0.05	8.2	0.05	28.8	0.05	26.5
Constant		−1.36	−6.5	−1.89	−5.0	0.04	0.0	1.12	4.1

The numbers in the headline are: 1=forward ($R^2 = 0.282$, 44 variables), 2=backward ($R^2 = 0.369$, 81 variables), 3=forward ($R^2 = 0.4415$, 215 variables), 4=backward ($R^2 = 0.4442$, 258 variables). The first two regressions (1 and 2) have clustered standard errors (each study is one cluster), the last two do not. The selection criteria is the significance of each variable in the first two regressions and the AIC improvement in the latter. c and t are the coefficients and the corresponding t-values of the included variables. The adjusted in- and exclude probabilities are 0.005 and 0.03 in the first two regressions. The column *var* refers to the variation of a variable (i.e., the percentage of valid observations); the maximum variation for dummy variables is fifty percent. The reference category for dummies is usually the opposite property or, in the case of multiple categories, the missing values.

end of the table 3.52

As expected, both backwards methods yield more variables (84 by the sw and 270 by stepAIC algorithm) than those methods which start with an empty set (44 and 234). Due to the unavailability of an existing implementation in R, the stepAIC-method does not use clustered standard errors and therefore yields even more significant variables. Out of the 270 (234) variables 203 (174) are significant at a 0.03 level and still 156 (141) are significant at a 0.005 level.

It is somewhat surprising that the stepAIC-algorithm has selected so many authors to be of major influence. While most of the coefficients have the same sign as in the bivariate comparison (table 3.36), some do not: Piquero has published significantly larger (i.e., less negative) normalized t-values but his dummy has a negative sign in regression three and four. This means that much of his larger (normalized) t-values can be explained by other factors. There are only two authors who are included in all four regressions: Ehrlich and Virén (the Virén-dummy and the Finnish author dummy are identical) and both are negative. All other authors are either in line with the bivariate results or appear only in one regression. Whether these author dummies should be interpreted as evidence of a publication bias or are more associated with unobserved heterogeneity is difficult to judge in this context. Among the nationality of the authors, Sweden, Switzerland and the Netherlands stand out, all bearing large positive coefficients, all being highly significant.

Concerning the discipline of the authors, experience suggests that criminologists and sociologists have larger (i.e., more positive) normalized t-values than economists. While this is supported by the bivariate analysis (table 3.27), the stepwise regressions show a more differentiated picture. Although the general trend is supported by the publication- and discipline-dummies, there are differences within each discipline. For example, studies published in "Criminology" have larger (normalized) t-values while those in "Criminal Law and Criminology" are smaller. The same can be observed for "Social Problems" and "Social Sciences Quarterly".

Among the implemented deterrence variables, the clearance rate seems to be very interesting. While it is associated with more negative (normalized) t-values in the bivariate case, its influence has been reversed in the regressions above. The mean sentence length has now a positive effect on the (normalized) t-values (i.e., is less in favor of the deterrence hypothesis) while it is ambiguous in the bivariate case. The same applies to the 'other" deterrence variables. The rest of the deterrence variables are in line with table 3.44.

Furthermore, it is interesting to note that those models which use transformed exogenous variables tend in different directions. Untransformed variables or those in logarithms are associated with more negative results. The opposite is the case for the endogenous variables. Most covariates have still the same influence as in table 3.45, but some change. Age and especially personal characteristics and the social class switch their signs and are now associated with more negative (normalized) t-values. Marital status and the time trend also switch, but in the opposite direction.

In table 3.40 it seems to be that studies using surveys or crime data are associated with larger (normalized) t-values. This picture is put into perspective by the results above. When other factors are accounted for (e.g., disciplines of authors or publishers), results based on surveys and crime data are even associated with more negative (normalized) t-values. As expected, the results for studies about the death penalty yield larger t-values.

Results based on misdemeanors are significantly related to more negative (normalized) t-values, while the opposite can be found for results which are based on accidents. This fits very well to the common expectation that offenses based on utility considerations are more readily accessible by deterrence measures, while drunk driving is often committed by people who do not react to deterrence. The latter is also supported by the inclusion and the positive sign of the drunk driving variable. Only vehicle theft does not fit into this picture. Regarding the studied offenses, some interesting observations emerge: although only based on very few observations, the deterrent effects based on the possession of hard drugs are strongly negative, while those based on the possession of soft drugs are almost equally strong but positive. This supports the view that possession (and usage) of marijuana, for example, is less affected by anti-drug laws, while people more readily react when, for example, crack is involved. Obviously, this might be partially explained by the more severe penalty for possessing hard drugs and by the larger public acceptance of soft drugs.

It is somewhat surprising that almost all included variables describing the method of analysis bear a negative sign. The correlation-dummy is a rare exception with inconsistent results (positive sign by the sw-algorithm, negative sign by the stepAIC-algorithm). Although COX-regressions are only rarely used, the associated dummy stands out because it is chosen by all four regressions. The opposite effect is found for 2SLS and GLS models which are significantly associated with smaller (i.e., more negative) estimates. Furthermore, it seems to be the case that methods which do not consider simultaneity overestimate deterrent effects.

Other noteworthy observations are that results from studies using Canadian data are less in favor of deterrence. The same applies when the nation under study does not belong to the most frequent nations. Results which are entered by the user tr into the data base appear to be significantly more negative. This is probably explained by the fact that he entered all economic studies while all other users entered the sociological and criminological studies; tr also worked at a different location, while all other users worked in the same department. The possibility of any intentional bias can be excluded. Positively signed offenses included in the regressions are drunk driving, environmental offenses, fraud, tax evasion, negligent assault, burglary, vehicle theft. Negatively signed are severe larceny, assault, drug related crimes, as well as assault. Results are more in favor of the deterrence hypothesis if deterrence is the focus of the study. The high significance of the realized sample sizes is a bit odd. These variables are not diversified by the sign of the results, as is the case with the number of observations. Since the latter are included in every regression, the relationship, explained in section 3.4, should already be taken care of. Again, also the stepwise regression methods include some variables with only very little variation. These seem to catch some oddities in the data which cannot be explained sufficiently by more general variables.

3.6.6 Bayesian Model Analysis

> *In a world in which the price of calculation continues to decrease rapidly, but the price of theorem proving continues to hold steady or increase, elementary economics indicates that we ought to spend a larger and larger fraction of our time on calculation.*
>
> *Wilder J. Tukey, American Statistican, 1986*

The basic idea for both approaches is quite intuitive: calculate the probabilities of all models using Bayes Theorem and chose the most probable (BMS) or use all of them (BMA), refer to Raftery et al. (1997) and Hoeting et al. (1999).

Bayesian Model Selection (BMS)

At first, for each model Δ, its probability over all possible models M_k, given the data D (δ_k is the vector of model parameters of Model M_k), is calculated:

$$P(\Delta|D) = \sum_{k=1}^{K} P(\Delta|M_k, D)P(M_k|D), \text{ with}$$
$$P(M_k|D) = \frac{(P(D|M_k)P(M))}{(\sum_{l=1}^{K} P(D|M_l)P(M_l))} \text{ and}$$
$$P(D|M_k) = \int P(D|\delta_k, M_k)P(\delta_k|M_k)d\delta_k.$$

Then, the models with the highest posterior probabilities are chosen. However, the implementation of the estimator poses several difficulties (also see Koop and Potter (2003) or Chipman et al. (2001)). The quality of the results may hinge on the selection of the hyper-parameters which are necessary for the calculations (Chipman et al., 2001). They can be chosen manually, calculated from the data or simply set to trivial values. Since we do not possess any usable information about the priors, we chose to use uninformative priors. As usually, there are several other problems to cope with: the huge model space involves all 2^N models. Therefore, optimization algorithms and monte carlo methods should be employed in our case. We resort to the BMA-package in R. Since the algorithm can only cope with 200 variables simultaneously and has no monte carlo features, we used a weakened version of test 4 from subsection 3.6.4 to preselect a reduced set of variables. The level of significance was chosen in such a way that 199 variables are included (plus the constant). Due to limited computational resources and algorithm restrictions we limited the number of variables, which are simultaneously included in each model, to a maximum of 50. Each BMA-regression then took about four days on a 4.2GHz Athlon X2 using only one core.

Basically, we will use BMS in a comparison of the methods in section 4.2 and simply select those variables with a posterior probability larger than 0.9 (which are essentially all variables used in the final BMA-model.

Bayesian Model Averaging (BMA)

BMA is essentially the same as BMS, with the exception of using the information of all considered models. Instead of the coefficient of the most probable model, BMA calculates the weighted (by their posterior model probability) average of each coefficient. Fernández et al. (2001b) compare BMA with EBA (as implemented by Sala-I-Martin (1997)) and find that BMA achieves better results. BMA has also been applied to crime data by Raftery et al. (1997); Fernández et al. (2001a); Liang et al. (2001) and Nott and Green (2004).

We note here that there are multiple other possibilities to chose the model-weights for calculating the coefficients. We chose BMA because it is used in the deterrence literature, is acknowledged by many researches and - certainly an important argument - is implemented in an available statistical software package. Hansen (2007) compares several averaging methods (based on the AIC, BIC, Mallows criterion and MMA - the Mallows Model Average estimator) and their performance. Performing a simulation he concludes that the MMA estimator has the lowest risk (expected squared error) and the weights based on the bayesian coefficients perform only better when the number of observations and R^2 is low. In contrast to the other weights, the risk of the BIC-based estimator is not decreasing in the number of observations. Overall, the MMA estimator is found to be performing best but could not be implemented into this analysis because the article was published too late. Nevertheless, the BMA-results are given in table 3.53.

Table 3.53: Multivariate analysis - bayesian model averaging

Variable	var	$p \neq 0$	coef.	sd
Study: size of first realized sample	58.8	100.0	−0.0001	0.000
Study: size of second realized sample	2.4	100.0	−0.0004	0.000
Study: tests of significance	9.4	100.0	−0.8635	0.119
Study: user, mw	1.4	100.0	0.8113	0.166
Study: author, Steven D. Levitt	1.7	100.0	1.2320	0.266
Study: author, Daniel S. Nagin	1.3	100.0	1.3570	0.281
Study: author, Isaac Ehrlich	1.0	100.0	−1.9980	0.327
Study: publication, criminology	20.5	100.0	0.4058	0.088
Study: publication, psychology	2.8	93.5	−0.7055	0.271
Study: institute, economics	41.8	100.0	−0.4757	0.088
Study: institute, miscellaneous	14.9	100.0	0.4654	0.101
Study: first population, Canada	4.3	100.0	0.9006	0.160
Study: first population, other country	6.2	100.0	0.59450	0.134
Study: sample unit, first population, states	21.9	100.0	0.5283	0.083
Study: sample individuals, first population, pupils	3.1	100.0	0.7591	0.187
Study: sample individuals, second population, population	2.6	100.0	0.8753	0.200
Study: complete sample	9.8	100.0	−0.8037	0.110
Study: main location > 500000 inhabitants	3.5	100.0	−1.6430	0.176
Study: does not check representativeness	26.9	100.0	0.4378	0.074
Study: mixed questions for pretest	2.1	100.0	1.1620	0.228
Estimate: deterrence is covariate	14.8	100.0	0.5168	0.094
Estimate: exogenous, index multiplicative	1.9	100.0	0.9222	0.230
Estimate: study type, death penalty	8.2	100.0	0.6196	0.122
Estimate: exogenous, crime data, arrest rate	9.5	100.0	−0.6628	0.115
Estimate: exogenous, crime data, conviction rate	3.9	100.0	−0.7776	0.168
Estimate: exogenous, crime data, police expenditures	4.0	100.0	0.7563	0.171
Estimate: exogenous, crime data, police strength	7.9	100.0	0.8597	0.128
Estimate: exogenous, crime data, probability dummy (regime shift)	3.5	100.0	−0.9804	0.191
Estimate: exogenous, crime data, severity dummy (regime shift)	3.6	100.0	−1.0500	0.183
Estimate: exogenous, survey, severity of punishment by others	0.4	100.0	2.1570	0.490
Estimate: exogenous, experiment, experimental variation of probability of detection	2.1	100.0	−1.9570	0.232
Estimate: exogenous, in differences	3.1	100.0	0.9154	0.207

continued on the next page…

. . . last page of table 3.53 continued

Variable	var	$p \neq 0$	coef.	sd
Estimate: endogenous, recidivism	0.7	94.7	1.3530	0.491
Estimate: endogenous, accidents	4.2	100.0	0.8311	0.181
Estimate: offense, drug possession (soft)	0.7	100.0	1.6650	0.408
Estimate: offense, drug possession (hard)	0.5	100.0	−2.5740	0.530
Estimate: offense, environmental crimes, violations of prescriptive limits	2.3	92.4	0.7232	0.292
Estimate: endogenous, binary category	9.6	100.0	−0.4611	0.116
Estimate: endogenous, not in logs	32.0	100.0	0.4240	0.077
Estimate: endogenous, other transformation	7.8	100.0	−0.9145	0.136
Estimate: covariate, fixed effects (spatial)	10.0	100.0	−0.66730	0.116
Estimate: covariate, poverty, welfare	6.4	100.0	0.78340	0.137
Estimate: covariate, urbanity	8.2	100.0	0.53860	0.123
Estimate: covariate, population (-growth)	11.5	100.0	0.43300	0.106
Estimate: no correction for simultaneity	19.3	100.0	−0.49180	0.088
Estimate: bivariate method, t-test for independent samples	0.5	100.0	−2.31400	0.427
Estimate: no test of significance	5.5	100.0	−0.65720	0.166
Estimate: square root of sample size for negative values	79.2	100.0	−0.01422	0.001
Estimate: square root of sample size for positive values	82.5	100.0	0.05388	0.002
constant		100.0	−0.96310	0.165

Bayesian model averaging with a maximum of 50 variables per regression. Algorithm supports only 200 variables, therefore 199 variables were preselected by EBA (weakened version of test D). The column *var* refers to the variation of a variable (i.e., the percentage of valid observations); the maximum variation for dummy variables is fifty percent. Properties of the best model: R^2 : 0.348, BIC: −2364; posterior probability: 0.805. The reference category for dummies is usually the opposite property or, in the case of multiple categories, the missing values.

end of the table 3.53

The results are in line with the previous results. Among the authors, Levitt, Nagin and Ehrlich are considered important enough to be included, while only the latter has a negative (finding more deterrent effects) impact. Studies published in a criminological journal have a positive effect, in opposition to psychological journals. When a Canadian population (or “other” country) is studied, the results are also more positive, while they are more negative when large cities are studied. When the death penalty is analyzed or the deterrence variable is just a covariate, the findings are less in favor of the deterrence hypothesis.

Among the deterrence variables, the arrest and conviction rate, as well as regime shifts and experimental variation of the detection probability are considered to be very important and have a negative effect. The influence of using police measures as deterrence variables keep their positive sign, while not correcting for simultaneity (which has negative effect) is also included. Drug possession, distinguished by soft and hard drugs, is also signed as expected.

The covariates considered to be most important are poverty, urbanity, population and the usage of spatial fixed effects, while only the latter has a negative influence. Studies which report tests of significance are associated with lower (more in favor of the deterrence hypothesis) normalized t-values while this is somewhat put into perspective on the estimate-level (having the opposite effect). As expected, the relationship between the square root of the number of observation and the t-values is considered to be of great importance.

3.6.7 Other Methodologies

What egotism, what stupid vanity, to suppose that a thing could not happen because you could not conceive it!

Philip Wylie and Edwin Balmer, When Worlds Collide, 1932

We also experimented with other methodologies and a much smaller test set. Although some interesting results could be extracted, we would have to recode the whole data set to be able to reasonably apply these techniques. Eventually, this would be beyond the scope of this study. Among these methods were *Decision Trees* and *Rough Set Data Analysis* (RSDA). While no clear cut results could be achieved from the Decision Trees - probably because our test set of 200 studies was too small - we could show promising results with RSDA (Rupp, 2005). For further information about RSDA refer to Pawlak (1982, 1991) and Pawlak (1999) and the references in our paper Rupp (2005).

Additionally, as indicated before, we used a trial version of PcGETS (Hendrey and Krolzig, 2000) to study our test set of 200 studies. Hoover and Perez (1999) show that the usability of PcGETS as a data mining tool depends on the data. With "good" data, an adjustment of the significance levels, such as in Lovell (1983), is not necessary in conjunction with PC-Gets. However, our data set seems to belong to the set of "bad" data because a preliminary analysis showed that the results were quite inconsistent and indecisive.

3.6.8 Interim Conclusion

If a person (a) is sick, (b) receives treatment intended to make him better, and (c) gets better, then no power of reasoning known to medical science can convince him that it may not have been the treatment that restored his health.

Peter B. Medawar, The Art of the Soluble, 1967

In the previous subsections we have found many important variables. However, having such a large arsenal of variables at our disposal makes the remark by Florax and de Groot (2002) noteworthy, "that almost any relationship can be shown to be either significantly positive or negative, provided the *correct* set of conditioning variables used".

All in all, it should be fairly obvious that a bivariate analysis, as conducted in section 3.5, is not sufficient to fully understand the relationships between the properties of the included studies and their results in regard to the deterrence hypothesis. Under the assumption that the methods of analysis are equally accurate, there are only a few variables which are judged to be significant and of the same sign by all methods. For example, all indicate that studies providing some sort of significance tests are significantly associated with results which are more in line with the deterrence hypothesis. The same applies when Issac Ehrlich is an author of a study, deterrence is the focus of a study, the studied offense is the possession of hard drugs, the endogenous variable is a dummy or when no correction for simultaneity is implemented. The opposite effect is found when a Canadian population is studied, the representativeness of the data is not checked, the death penalty is studied or when poverty, welfare, urbanity or population (-growth) are used as covariates. When a variable is included in every table - i.e., the variables just mentioned - all methods agree on their signs. However, when drawing conclusions, it might be more practical to weigh the methods differently or to consider variables which are not included in all tables.

The task will be to find those variables which are really meaningful. To accomplish this, two steps seem to be important. First, as already done above, to identify those variables which share the same properties in all estimations. Second, to judge the quality - in regard to precision and fit - of the estimators and to interpret the results accordingly. The latter is be done in chapter 4.

4 Assessing the Quality of the Results

It doesn't do to leave a live Dragon out of your calculations, if you live near him.

John R. R. Tolkien, The Hobbit, 1937

The combination of some data and an aching desire for an answer does not ensure that a reasonable answer can be extracted from a given body of data.

Wilder J. Tukey, American Statistician, 1986

In order to interpret the results from the previously performed regressions we have to answer two questions. First, whether we are interested in a good fit of the data or if we want a good prediction of a study's results. Second, which criteria are applicable to judge whether a fit or prediction is superior to another? A good fit requires a model which can replicate the data as well as possible. This means that it has to incorporate the main factors, as well as to handle specialities and oddities found in the data. Usually, this will lead to a model with a large pool of variables. On the other hand, predicting the results of an unknown study (i.e., a study that is not included in the estimation of the model) requires a model which catches the main influences while not contaminating its results by anomalies from known studies. This should lead to a model with a smaller pool of variables.

Since both aims - fitting and predicting - have their merits, we will analyze our models from section 3.6 under both points of view. A model with a good fit can be used to understand the importance and the effects of variables included in the model. Such models may help to understand the influence of very special variables (e.g., a specific author). On the other hand, a model which produces good predictions can be used to draw conclusions about unknown, future or hypothetic studies. It may also catch more general influences, since it should better model the general underlying mechanics. Combining both strategies - fitting and predicting - could also be helpful in putting any conclusions on a firmer ground than relying only on one method.

4.1 Modus Operandi

Do not go where the path may lead; go instead where there is no path and leave a trail.

Ralph W. Emerson

We use a large arsenal of criteria to judge how well a model performs. These criteria are split into two parts. First, we use various loss functions (described in subsection 4.1.1) to measure the difference between the estimated and actual values. These functions almost all relate to the error terms of each model and help us to identify important properties of a model's performance. Second, we calculate several classification rates (see subsection 4.1.2). These describe how well the endogenous variable is estimated by the models in respect to some more general and very important categories. The reasoning is that it may be more important to replicate, for example, a significant and negative (normalized) t-value than to approximate its actual value.

To study the quality of the various estimators we employ several approaches. To study the fit of the models we use the whole data set to calculate the fitted values. We also use some bootstrapping to get the first and second moments of the values of the loss functions. To predict the data, and for the bootstrapping, we randomly partition the data into a training and a test-set. The same estimators as described in subsections 3.6.3, 3.6.4, 3.6.5 and 3.6.6 are studied.

We chose to use two ways to construct the test- and training sets. First, we randomly chose 50% of the data to belong to the training set and the remaining 50% is assigned to the test set. Although the size of 50% is unusually large for a test set, it is prerogative in our case because these values are most likely not independent (many studies will be present in both sets). Second, we chose to assign 90% of all studies to the training set. The latter approach better corresponds to the idea of forecasting the outcome of an unknown (i.e., not included) study, while the first relates more to the prediction of random data in general. Each forecast or fit is calculated with ten independently and randomly generated training sets. The ten sets are the same for each estimator within the same approach (e.g., a training set of 50% or 90%).

4.1.1 Loss-Functions

Just because you have a choice, it doesn't mean that any of them has to be right.

Norton Juster, The Phantom Tollbooth, 1961

We employ a wide variety of loss functions to distinguish various characteristics of the estimators. A summary of fifty loss functions, from which we have taken some, is given by Andres and Spiwoks (2000). Each loss function has its own merits and justification. Most of them are symmetrical (punishing deviations in both directions equally), while asymmetrical loss functions are also possible but are, at least to some extent, quite arbitrary.

- RMSE: the root mean squared error $\sqrt{\frac{\sum_{i=1}^{N}(y_i-\hat{y}_i)^2}{N}}$.

  The most commonly used loss function. The lower its value, the better the estimates are. Small errors (< 1) are less important than larger (> 1) errors.

- Cor.: the Pearson correlation between y and $\hat{y}$. There shouldn't be any negative values; the closer to one the better the estimates are.

- Adj. R^2: the classic adjusted R squared, $1-(1-R^2)\frac{N-1}{N-k-1}$, with k being the number of regressors. It is the same as the correlation but adjusted for the ratio between the of number of regressors and observations.

- U: Theil's (new) inequality coefficient $\sqrt{\frac{\sum_{i=1}^{N}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{N}y_i^2}}$ and its decomposition
  - U.bias $=\frac{(\bar{y}-\bar{\hat{y}})^2}{\frac{1}{N}\sum_{i=1}^{N}(y_i-\hat{y}_i)^2}$,
  - U.var $=\frac{(s_y-s_{\hat{y}})^2}{\frac{1}{N}\sum_{i=1}^{N}(y_i-\hat{y}_i)^2}$,
  - U.cov $=\frac{2(1-\rho)s_y s_{\hat{y}}}{\frac{1}{N}\sum_{i=1}^{N}(y_i-\hat{y}_i)^2}$.

  U should be zero in the case of a perfect, and one for the most naive estimator (the constant; any values above one indicate that the estimator is worse than the naive estimator). The estimation errors are divided into U.bias (systematic error in the mean value), U.var (systematic error in the variance) and U.cov (unsystematic random error). These should add up to one (except for rounding errors). The perfect estimator has a U.cov of one.

- RMSPE: the root mean squared proportional error, $\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\frac{y_i-\hat{y}_i}{y_i})^2}$.

  It is similar to the RMSE but measures the error relative to the true values. Therefore, it is independent of scaling.

- CI.hit: the fraction of predicted values in a $c\cdot s_y$ interval of y. In our case we set c to 0.5. This measures whether or not the predicted values are "near" the actual values but does not consider the size of the errors (similar as in Koop and Potter (2003) but with a tighter bandwidth).

- Sign.: The percentage of correctly classified significant (normalized) t-values at a 5% level (the categories are negative and significant as well as positive and significant). Although this measure does belong to the classification ratings, it is also included in the loss functions because a similar measure (the percentage of negative significant (normalized) t-values) is included in many of the tables in section 3.5 as well.

- Neg2pos4: a loss function which punishes large deviations much harder in the case of positive values, $\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(1_{[y_i\leq 0]}(y_i-\hat{y}_i)^2+1_{[y_i>0]}(y_i-\hat{y}_i)^4\right)}$

  We implement this loss function because we have an abundance of negative but only relatively few positive values. With this function we see whether an estimator fares better or worse with positive values.

- FsRMSE: false sign root mean squared error, $\sqrt{\frac{1}{N}\sum_{i=1}^{N}1_{[y_i\cdot\hat{y}_i<0]}(y_i-\hat{y}_i)^2}$.

  This function is very similar to the RMSE-function but only punishes those estimates which carry the wrong sign. We implemented this because the sign of an estimate is, to some degree, even more important than the extent of its deviation.

- Min. dev. and Max. dev.: the maximum $\max(y,\hat{y})$ and minimum $\min(y,\hat{y})$ deviation. Since there are some outliers in the data which are not catched by any model, the minimum and maximum deviation are only of academic importance.
- Mean pos.: the mean positive deviation $\frac{1}{N}\sum_{i=1}^{N} 1_{[\hat{y}_1>y_i]} \cdot |y_i-\hat{y}_i|$,
- Mean neg.: the mean negative deviation $\frac{1}{N}\sum_{i=1}^{N} 1_{[\hat{y}_1<y_i]} \cdot |y_i-\hat{y}_i|$,
- Mean abs.: the mean absolute deviation $\frac{1}{N}\sum_{i=1}^{N} |y_i-\hat{y}_i|$.

  By comparing the mean absolute and the mean squared error, we can judge wether the estimator tends to vary around the actual values more closely with some large errors (larger RMSE and smaller mean absolute error) or has less large errors but deviates more in general (smaller RMSE and larger mean absolute error).
- A Log-predictive-Score LPS, somewhat similar to the original (see Good (1952)), but defined as $-\frac{1}{N}\sum_{i=1}^{N} \log 2(1-\mathcal{N}(|\hat{y}_i-y_i|/\hat{s}))$; $\mathcal{N}$ is the inverse CDF of a standard normal distributed random variable. A perfect fit would result in an LPS of 0.
- An adjusted LPS by adding a penalty $2k/N$ for the number of used regressors k (similar as for the AIC).

Additionally, we perform a general encompassing test (a bit more simplified than in Clements and Harvey (2004)), by regressing $y=\sum_{i=1}^{\#E}\beta_i\hat{y}^{(i)}+\varepsilon$ and analyzing the calculated coefficients cEncomp. and the respective p-values pEncomp. ($\#E$ is the number of competing estimators $\hat{y}^{(s)}$). Good estimators should have coefficients near one and low p-values because they should contain most of the required information (in the sense of minimizing the sum of squared errors).

When possible, we also calculate the 95% confidence intervals of the mean of each loss function to see whether any method is significantly superior to the naive estimators. We call a model A superior (inferior) to model B in regard to a certain loss function f, if the confidence interval of f for model A does not include the mean value of f for model B and the mean of $f(A)$ is better (worse) than the mean of $f(B)$.

4.1.2 Classification Ratings

While the loss-functions reflect the behavior of the error terms, it is also important to know whether the estimators are able to catch more general characteristics and behavior of the data. For example, it might be more important to estimate a t-value correctly to be negative or negative and significant, than whether it is -2.5 or -5.5. To study these characteristics we employ several categories:

Sign All (normalized) t-values are categorized to be negative or positive. We use this to have a look at the general direction of the estimates. The correct sign might be considered more important than an exact estimate. Moreover, we examine each sign separately to see whether an estimator fares better with results which tend to approve or disapprove of the deterrence hypothesis.

Pos. Sign A (normalized) t-values belongs to this category if it is significant and positive.

Neg. Sign A (normalized) t-values belongs to this category if it is significant and negative.

20% sign. A (normalized) t-value belongs to this category, if it is significant at a 20% level (i.e., $|t| > 1.28$). Since values around zero might be considered as noise, we ignore these results in this category.

20% pos. A (normalized) t-value belongs to this category, if it is significant at a 20% level and positive (i.e., $t > 1.28$).

20% neg. A (normalized) t-value belongs to this category, if it is significant at a 20% level and negative (i.e., $t < -1.28$).

5% sign. A (normalized) t-value belongs to this category, if it is significant at a 5% level (i.e., $|t| > 1.96$). Naturally, it is very interesting to see how well the estimators handle the significant (normalized) t-values. We use the 5% level to discard results which do not significantly approve (disapprove) of the deterrence hypothesis.

5% pos. A (normalized) t-value belongs to this category, if it is significant at a 5% level and positive (i.e., $t > 1.96$).

5% neg. A (normalized) t-value belongs to this category, if it is significant at a 5% level and negative (i.e., $t < -1.96$).

For these categories, which are assumed to be non-empty sets, we calculate two measures:

Precision The percentage of values which are correctly classified (e.g., if a category actually contains n values, and of these $\hat{n}$ are not estimated correctly, the precision is $n - \hat{n}$ divided by n). In the perfect case the precision is one, and zero in the worst case.

Error The error rate (if m values are estimated to belong to a category, but of these $\hat{m}$ are not correct, the error rate is $\hat{m}$ divided by m). In the perfect case the error rate is zero, and one in the worst case. In the case that $m = 0$ the error rate is defined to be zero.

Total miss rate The sum of not correctly and falsely classified values divided by the total number of values belonging to that category (e.g., $[(1 - \text{precision})n + (\text{error rate})m]$ divided by n). In the perfect case the total miss rate is zero, and limited by $\hat{m}$ in the worst case.

As in subsection 4.1.1, we calculate the 95% confidence intervals of the mean of each classification rate to see whether any method is significantly superior (inferior) to the naive estimators.

4.2 The Tournament

A theory is a good theory if it satisfies two requirements: it must accurately describe a large class of observations on the basis of a model that contains only a few arbitrary elements, and it must make definite predictions about the results of future observations.

Stephen W. Hawking, A Brief History of Time, 1988

To test the reliability of the calculated estimators, we let them compete against each other in a tournament of estimation quality. Each estimator has to proof itself in its precision and classification quality.

4.2.1 The Competitors

There is no need for these hypotheses to be true, or even to be at all like the truth; rather one thing is sufficient for them - that they should yield calculations which agree with the observations.

Andreas Osiander, Preface to Copernicus' De Revolutionibus, *1543*

We gathered all estimators (from subsections 3.6.3, 3.6.4, 3.6.5 and 3.6.6) and assigned them into different groups:

naive The naive approaches which select either none or all variables.

- **SET0** The most naive estimator. In case of the loss-function it is the mean weighted (normalized) t-value of all observations (i.e., a regression with the constant only), while it is a random guess[1] for the classification ratings.
- **SET1** The OLS regression with all 492 variables (some variables are dropped due to singularity problems).

SET2 All remaining variables after removing those which are responsible for singularity problems and which have a p-value > 0.1; 77 variables remained. Although this could also be named naive, we take this specification as the most simple model.

EBA Extreme Bounds Analysis with different inclusion criteria. All variables with a stability-coefficient which does not lie within a 95% CI of the mean stability-coefficient are excluded beforehand[2].

[1] The category of a random (normalized) t-value estimated by the mean weighted (normalized) t-value plus a normal distributed random variable (the standard deviation is that of the full data set without the 1%-quantiles).

[2] The stability-coefficient is calculated as the quotient of the mean and median (normalized) t-value. All variables which lie outside an interval of twice its standard deviation around the mean stability-coefficient are excluded because these results seem to be unreliable.

SET3 Extreme Bounds Analysis with the CDF test (A); 254 variables are selected.

SET4 Extreme Bounds Analysis with the strong sign test (B); 156 variables are selected.

SET5 Extreme Bounds Analysis with the extreme CDF test (C); 119 variables are selected.

SET6 Extreme Bounds Analysis with the absurd test (D); 80 variables are selected.

stepwise Stepwise (backward and forward) regressions.

SET7 Forward stepwise regression (starting with no variables) based on the significance of each variable; 43 are variables selected.

SET8 Backward stepwise regression (starting with all variables) based on the significance of each variable; 80 are variables selected.

SET9 Backward stepwise regression (starting with no variables) based on the AIC improvement; 269 variables are selected.

SET10 Forward stepwise regression (starting with all variables) based on the AIC improvement; 233 variables are selected.

BMA Bayesian Model Selection and Averaging. Due to computational limitations only 50 variables are allowed to be included in any submodel at any time. Eventually, the best 15 models are selected.

SET11 All variables are selected with a posterior probability ≥ 0.9 after a Bayesian Model Averaging procedure; 49 variables are selected.

SET12 Full Bayesian Model Averaging in every run. 49 variables are selected (the constant is the 50th).

4.2.2 The Contest

It is far better to foresee even without certainty than not to foresee at all.

Jules H. Poincaré

In the following we present and compare the performance of the different estimators. The first part, section 4.2.2, contains the measures to assess the quality of the estimators. They describe how well the estimators predict and fit the actual data. The second part, section 4.2.2, shows the percentages of correctly and incorrectly classified (normalized) t-values. These are used to judge the performance of the estimators in reproducing certain properties of the (normalized) t-values.

Predicting and Fitting Performance

Tables 4.1 and 4.2 contain the resulting values of the loss-functions for the 50%-observation based test sets, while tables 4.3 and 4.4 contain those which are based on the 10%-study based test sets. The rows contain the loss-functions, while the various competing methods are located in the columns. To simplify the interpretation, the number of variables used in each method is given in the second row. Bold cells contain the best value of all methods (in the case of the error statistics this means more similarity to the normal distribution). Light (dark) cells indicate that the model is superior (inferior), as described in subsection 4.1.1, to the best naive model (SET0 or SET1).

It is obvious that the SETS 7, 8 and 11 perform very well compared to all other methods. They perform best in many criteria and especially the backward stepwise regression (based on significance) has a high correlation, the best root mean squared error (for all observations as well as for those with a false predicted sign), the best mean absolute error and log predictive score. Furthermore, the variance of its residuals is the smallest and the encompassing test does also state that it does contain relevant information. Considering the fact that the stepwise forward method just uses 43 variables it performs very well. This is reflected by the best adjusted LPS (but not the adjusted R^2), the hit ratio and the relative RMSE. The same applies to the BMS-set, which has the highest correlation and, although not the best, many fairly good values. It should be mentioned that the means given for BMA inhibit a very large deviation - in some runs it performed very well, in others extraordinarily bad.

In table 4.2 we test how well the estimators are in reproducing the data. The whole data set is used to establish the estimator, and a randomly chosen 50% set of the data is then estimated. This is repeated ten times. The same notation as in table 4.3 applies. This is done to study whether and by how much any method performs significantly better or worse than the naive methods[3], when the estimators are based on the whole data set.

It is not surprising that SET1 has the best RMSE, correlation values and is the only method chosen by the encompassing test - and that its residuals are "the best" - because the model is optimized according to these criteria. Nevertheless, the stepwise estimators perform fairly well: the relative RMSE is the best for SET7 while SET9 and SET10 are the only estimators which are not significantly worse than the full OLS estimator and even outperform it in some criteria (Sign. and fsRMSE). As could be expected, estimators which are based on many variables perform better than those based on only a small set of variables in fitting the data.

Instead of using very large test sets (50% of the data) and to lessen the effect that many observations will not be independent (because they belong to the same study), we repeat the procedure with a 90% vs. 10% partition on the study-level. Thus, the (normalized) t-values belonging to a randomly chosen 90% set of all studies are assigned to the training set, while the rest remains in the test set. Due to time constraints and the mediocre performance of BMA, we removed the bayesian model averaging estimator from the sets and did not recalculate the set of variables for the bayesian model selection approach for each run[4].

[3] The OLS estimator with all variables and data shows the best performance in many criteria (eg. RMSE, correlation, encompassing test) because it is constructed that way.

[4] We use the same set of variables derived from the whole data set and merely recalculated the coefficients for each training set.

Table 4.1: How well the models predict random data sets

Method # var.	SET0 0	SET1 492	SET2 77	SET3 254	SET4 156	SET5 119	SET6 80	SET7 43	SET8 80	SET9 269	SET10 233	SET11 49	SET12 49
RMSE	3.076	6.557	2.803	2.928	2.813	2.775	2.763	4.025	**2.667**	4.212	2.742	2.724	6.478
Cor.	0.000	0.295	0.444	0.431	0.464	0.475	0.475	0.377	**0.520**	0.443	0.520	0.491	0.333
Adj. R^2	-0.000	0.169	0.431	0.383	0.437	0.455	0.461	0.368	**0.508**	0.393	0.483	0.483	0.322
U	0.907	1.937	0.828	0.866	0.832	0.820	0.817	1.213	**0.788**	1.244	0.811	0.805	1.914
U.bias	0.003	0.004	0.003	**0.003**	0.003	0.003	0.003	0.003	0.003	0.004	0.004	0.003	0.149
U.var.	0.997	0.109	0.219	0.101	0.147	0.166	0.187	0.292	0.191	0.132	**0.095**	0.197	0.186
U.cov.	0.000	0.887	0.778	0.897	0.851	0.830	0.810	0.705	0.806	0.865	**0.902**	0.801	0.664
RMSPE	18.562	27.128	17.988	25.230	22.280	21.650	19.268	**16.556**	21.319	27.479	26.334	19.757	48.125
CI.hit	0.556	0.479	0.595	0.541	0.570	0.575	0.578	**0.617**	0.600	0.557	0.577	0.591	0.424
Sign.	0.000	0.494	0.397	0.475	0.463	0.457	0.426	0.471	0.487	**0.516**	0.513	0.449	0.349
Neg2pos4	16.509	1078.900	13.460	12.594	11.616	11.450	11.363	158.39	11.368	277.933	**11.010**	11.375	620.187
FsRMSE	12.774	149.470	8.073	10.590	8.544	8.092	7.822	65.584	**6.903**	52.646	9.149	7.570	30.235
Min. Dev.	-18.632	-79.830	-19.350	-19.507	-18.896	-18.643	-18.662	-20.400	-18.363	-47.543	-19.650	-18.690	**-17.611**
Max. Dev.	54.694	50.853	51.002	50.172	50.061	50.082	50.057	68.295	50.371	49.716	**49.554**	50.353	75.909
Mean pos.	2.014	2.591	1.869	2.054	1.954	1.916	1.909	2.061	**1.833**	1.996	1.904	1.855	4.157
Mean neg.	-1.906	-2.511	-1.658	-1.813	-1.704	-1.684	-1.666	-1.568	**-1.564**	-1.912	-1.651	-1.629	-2.012
Mean abs.	1.963	2.552	1.765	1.932	1.829	1.600	1.787	1.804	**1.698**	1.954	1.778	1.742	3.891
LPS	0.830	1.056	0.731	0.794	0.742	0.722	0.722	0.713	**0.682**	0.764	0.718	0.696	1.968
Adj. LPS	0.830	1.132	0.742	0.833	0.766	0.740	0.734	0.720	**0.695**	0.805	0.753	0.704	1.975
cEncomp.	0.000	-0.165	-0.060	-0.249	0.139	0.027	0.110	0.119	0.350	0.226	0.445	0.223	-0.295
pEncomp.		0.000	0.102	0.015	0.359	0.430	0.323	0.251	**0.000**	0.048	0.000	0.103	0.083
Mean	**-0.003**	-0.214	-0.068	-0.084	-0.085	-0.080	-0.063	-0.059	-0.087	-0.123	-0.109	-0.063	-2.314
Median	-3.632	-5.343	-3.549	-4.048	-3.850	-3.745	-3.705	-3.432	**-3.391**	-3.901	-3.666	-3.561	-8.410
Sd	2.236	2.834	2.067	2.263	2.146	2.095	2.080	2.011	**1.963**	2.159	2.063	2.018	3.508
Skewness	-0.509	-0.455	-0.462	-0.386	-0.445	-0.439	-0.425	-0.550	-0.458	**-0.351**	-0.388	-0.400	-0.665
Kurtosis	4.542	4.321	4.588	3.771	4.039	4.020	4.062	5.002	4.201	**3.685**	3.839	4.024	5.438

10 runs. Training sets contain random 50% of the data (rest belongs to the test sets). *SET0*: no variables. *SET1*: all variables. *SET2*: all significant variables (< 0.1) from SET1. *SET3-6*: EBA with criterion A-D. *SET7/8*: stepwise forward/backward (based on significance). *SET9/10*: stepwise backward/forward (based on AIC improvement). *SET11*: BMS (maximum of 49 variables at once). *SET12*: BMA (maximum of 49 variables at once). *# var.* is the number of variables used by each method. Light/dark grey cells: 95%-CI does not contain the best average of the naive approaches (i.e., is better/worse than the best of SET0/1). Bold cells are the best of each row. Rows below the line are statistics of the error term (excluding the smallest and largest 1% of the errors) and only the best (i.e., similarity towards $\mathcal{N}(0.\sigma^2)$) cells are marked.

Table 4.2: How well the models fit random data sets

Method	SET0	SET1	SET2	SET3	SET4	SET5	SET6	SET7	SET8	SET9	SET10	SET11	SET12
# var.	0	492	77	254	156	119	80	43	80	269	233	49	49
RMSE	3.073	**2.228**	2.560	2.391	2.442	2.455	2.497	2.611	2.444	2.248	2.257	2.491	2.491
Cor.	0.000	**0.689**	0.552	0.628	0.607	0.601	0.583	0.527	0.606	0.681	0.678	0.585	0.585
Adj. R^2	-0.000	0.633	0.541	0.597	0.587	0.586	0.572	0.521	0.596	0.653	**0.653**	0.579	0.579
U	0.906	**0.658**	0.756	0.706	0.722	0.725	0.738	0.771	0.722	0.664	0.667	0.736	0.736
U.bias	0.001	**0.000**	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.000	0.000	0.001	0.001
U.var.	1.000	**0.176**	0.284	0.222	0.235	0.241	0.252	0.301	0.238	0.181	0.184	0.251	0.252
U.cov.	0.000	**0.824**	0.716	0.778	0.765	0.759	0.748	0.699	0.762	0.819	0.817	0.748	0.748
RMSPE	19.016	21.705	17.168	22.091	20.188	20.826	18.273	**16.526**	20.714	23.205	23.627	18.224	18.326
CI.hit	0.557	**0.664**	0.622	0.623	0.631	0.623	0.610	0.629	0.624	0.664	0.655	0.608	0.608
Sign.	0.000	0.567	0.433	0.553	0.522	0.515	0.455	0.483	0.511	**0.584**	0.584	0.467	0.469
Neg2pos4	16.588	**8.835**	12.879	9.884	10.105	10.119	10.234	12.786	10.800	9.033	9.282	10.372	10.380
FsRMSE	12.947	6.265	6.564	6.128	6.209	6.290	6.463	8.337	6.745	**6.099**	6.190	6.429	6.383
Min. Dev.	-18.664	-16.042	-17.481	-15.739	-15.760	**-15.676**	-15.830	-16.648	-16.150	-16.166	-16.011	-15.697	-15.698
Max. Dev.	54.663	**48.579**	50.551	49.250	49.082	48.994	49.186	49.734	49.757	48.821	48.704	49.695	49.691
Mean pos.	2.002	**1.488**	1.735	1.627	1.673	1.670	1.740	1.807	1.687	1.489	1.507	1.727	1.724
Mean neg.	-1.919	**-1.376**	-1.555	-1.506	-1.535	-1.559	-1.563	-1.499	-1.490	-1.410	-1.403	-1.566	-1.568
Mean abs.	1.961	**1.431**	1.642	1.565	1.603	1.614	1.649	1.642	1.585	1.448	1.454	1.644	1.644
LPS	0.829	**0.564**	0.684	0.629	0.649	0.654	0.672	0.696	0.645	0.572	0.575	0.669	0.669
Adj. LPS	0.829	0.639	0.708	0.668	0.672	0.673	0.684	0.702	0.657	0.613	**0.611**	0.677	0.677
cEncomp.	0.000	0.901	0.010	0.063	-0.054	0.052	-0.062	-0.025	0.017	-0.000	0.089	1.406	-1.418
pEncomp.		**0.000**	0.293	0.504	0.411	0.574	0.402	0.558	0.426	0.626	0.484	0.477	0.480
Mean	0.029	-0.014	**-0.001**	-0.008	-0.001	-0.005	-0.006	-0.008	-0.019	-0.019	-0.015	-0.009	-0.005
Median	-3.601	**-2.760**	-3.313	-3.227	-3.226	-3.227	-3.308	-3.292	-3.080	-2.836	-2.831	-3.294	-3.283
Sd	2.236	**1.660**	1.914	1.823	1.864	1.878	1.915	1.935	1.836	1.678	1.684	1.908	1.911
Skewness	-0.507	-0.312	-0.463	-0.413	-0.449	-0.427	-0.474	-0.543	-0.453	**-0.301**	-0.347	-0.381	-0.364
Kurtosis	4.543	3.978	4.387	4.026	4.158	4.139	4.139	4.805	4.174	**3.934**	3.977	3.917	3.939

10 runs. Training sets contain random 50% of the data (rest belongs to the test sets). *SET0*: no variables. *SET1*: all variables. *SET2*: all significant variables (< 0.1) from SET1. *SET3-6*: EBA with criterion A-D. *SET7/8*: stepwise forward/backward (based on significance). *SET9/10*: stepwise backward/forward (based on AIC improvement). *SET11*: BMS (maximum of 49 variables at once). *SET12*: BMA (maximum of 49 variables at once). *# var.* is the number of variables used by each method. Light/dark grey cells: 95%-CI does not contain the best average of the naive approaches (i.e., is better/worse than the best of SET0/1). Bold cells are the best of each row. Rows below the line are statistics of the error term (excluding the smallest and largest 1% of the errors) and only the best (i.e., similarity towards $\mathcal{N}(0,\sigma^2)$) cells are marked.

Table 4.3: How well the models predict random studies

Method	SET0	SET1	SET2	SET3	SET4	SET5	SET6	SET7	SET8	SET9	SET10	SET11
# var.	0	492	77	254	156	119	80	43	80	269	233	49
RMSE	2.867	4.024	2.590	2.852	2.609	2.542	2.510	2.504	**2.412**	2.584	2.510	2.436
Cor.	0.000	0.205	0.466	0.396	0.476	0.502	0.506	0.511	**0.558**	0.510	0.534	0.543
Adj. R^2	-0.000	0.064	0.453	0.345	0.449	0.483	0.493	0.505	**0.546**	0.466	0.499	0.535
U	0.882	1.263	0.797	0.886	0.805	0.783	0.773	0.767	**0.741**	0.800	0.775	0.748
U.bias	0.006	0.008	0.011	0.008	0.008	0.008	0.008	0.008	0.008	0.008	**0.005**	0.010
U.var.	0.996	**0.048**	0.229	0.086	0.137	0.172	0.198	0.297	0.187	0.072	0.093	0.223
U.cov.	0.000	**0.945**	0.762	0.908	0.856	0.821	0.796	0.697	0.806	0.921	0.904	0.769
RMSPE	15.763	35.518	16.151	20.368	19.391	20.154	16.317	**14.036**	19.931	25.841	24.808	17.368
CI.hit	0.514	0.332	0.557	0.476	0.516	0.530	0.531	**0.593**	0.565	0.491	0.502	0.567
Sign.	0.000	0.457	0.414	0.483	0.502	0.501	0.424	0.492	0.516	**0.521**	0.515	0.464
Neg2pos4	12.558	22.606	11.636	11.281	10.320	10.216	9.941	10.558	10.395	10.767	10.618	**9.640**
FsRMSE	11.509	26.665	8.788	11.561	9.308	9.028	8.891	8.888	8.544	10.261	10.203	**7.356**
Min. Dev.	-11.599	-14.217	-11.136	-11.541	-11.150	-11.185	-10.884	-10.957	-11.106	-11.282	-10.943	**-10.624**
Max. Dev.	17.879	17.623	15.766	15.831	15.233	15.234	15.160	15.790	15.053	**14.407**	14.525	15.341
Mean pos.	2.001	2.989	1.843	2.155	1.948	1.887	1.838	1.818	**1.744**	1.947	1.880	1.803
Mean neg.	-1.841	-2.844	-1.624	-1.873	-1.697	-1.639	-1.630	**-1.436**	-1.508	-1.755	-1.679	-1.515
Mean abs.	1.914	2.920	1.731	2.012	1.820	1.760	1.733	**1.621**	1.622	1.855	1.778	1.656
LPS	0.928	1.629	0.801	0.959	0.830	0.796	0.781	0.746	**0.726**	0.835	0.795	0.739
Adj. LPS	0.928	1.705	0.813	0.998	0.854	0.815	0.794	0.753	**0.738**	0.876	0.831	0.747
cEncomp.	0.000	-0.354	-0.072	-0.457	0.106	-0.032	0.170	0.239	0.198	0.486	0.510	0.219
pEncomp.		**0.000**	0.355	0.015	0.461	0.372	0.301	0.197	0.243	0.163	0.080	0.313
Mean	0.002	**-0.020**	-0.012	-0.092	-0.064	-0.070	-0.071	-0.121	-0.055	-0.107	-0.081	-0.090
Median	-3.272	-6.146	-3.415	-4.092	-3.742	-3.596	-3.421	-3.391	**-3.212**	-3.814	-3.680	-3.370
Sd	2.161	3.478	1.997	2.314	2.100	2.023	2.000	1.919	**1.891**	2.109	2.041	1.926
Skewness	-0.589	**0.006**	-0.538	-0.418	-0.430	-0.476	-0.492	-0.769	-0.527	-0.269	-0.344	-0.548
Kurtosis	5.299	4.123	4.602	3.673	3.825	3.944	4.135	5.789	4.557	**3.149**	3.423	4.261

10 runs. Training sets contain random 90% of all studies (all other studies belong to the test sets). *SET0*: no variables. *SET1*: all variables. *SET2*: all significant variables (< 0.1) from SET1. *SET3-6*: EBA with criterion A-D. *SET7/8*: stepwise forward/backward (based on significance). *SET9/10*: stepwise backward/forward (based on AIC improvement). *SET11*: BMS (maximum of 49 variables at once). *# var.* is the number of variables used by each method. Light/dark grey cells: 95%-CI does not contain the best average of the naive approaches (i.e., is better/worse than the best of SET0/1). Bold cells are the best of each row. Rows below the line are statistics of the error term (excluding the smallest and largest 1% of the errors) and only the best (i.e., similarity towards $\mathcal{N}(0,\sigma^2)$) cells are marked.

Compared with table 4.1 the results in table 4.3 are quite similar. However, the good performance of SET8, the backward stepwise regression, is not as dominating as before. Other stepwise procedures (SET7 and SET9) get closer to SET8. Although the variables for the BMS approach are not recalculated, SET11 performs much better.

Table 4.4 is the analogy to table 4.2. The estimators are based on the full data set and their performance in fitting random 10% sets of the studies are compared.

There are not many differences between the results from the test sets based on random 10% studies or 50% data. Of course, SET1 is the best while the StepAIC results come second.

Classification Performance

For reasons of parsimony, the classification ratings are given only for the 10%-test sets, which can be interpreted as a simulation of the estimation of unknown studies. Tables 4.5 and 4.6 report the statistics of the precision of the estimates for predicting and fitting unknown studies. Again, the columns contain the models[5] while the rows display the categories. SET1 is not a constant anymore (which does not make much sense when classifying observations) but random draws from the statistical distribution[6] of the (normalized) t-values. The last column (N) contains the number of observations in the various categories. To simplify the interpretation, the number of variables used in each method is given in the second row. Finally, the cells report the precision ratings (the number of correctly classified observations divided by N). Again, bold cells contain the best value of all methods. Light (dark) cells indicate that the model is superior (inferior) to the best naive model (all or no variables used).

Overall, SET7 (which uses only 43 variables) performs best when classifying negative (normalized) t-values, while the AIC-based stepwise estimators are somewhat better in predicting positive values. The naive approach (SET1) perform worse in almost all categories (except the positive (normalized) t-values which are significant at a 5% level). Random guessing is especially bad.

When all studies are used to construct the estimators, the picture changes somewhat. Although the naive approach is still much better in fitting than in predicting, it performs not as well. SET1 is only the best in one category while SET9 and SET10 are not significantly worse in any category. SET9 seems to perform best in this case.

Tables 4.7 and 4.8 contain the classification error statistics of the models of predicting and fitting other studies. The models are in the columns and the statistics are given in the rows (the second row reports the number of variables used in each model). The rows are organized in groups of three lines and contain the average values of all ten runs.

The first line reports the category and the error rate, which is calculated by the percentage of falsely classified estimates in that category. This value is the number of estimates which actually

[5] Since the results by BMA are unreliable (very high variance in their quality) and would have taken several weeks to compute, we have, again, omitted BMA from the study-based analysis and did not recalculate the sets for BMS.
[6] In fact, we did not draw from the sample but used the normal distribution with the corresponding moments.

Table 4.4: How well the models fit random studies

Method	SET0	SET1	SET2	SET3	SET4	SET5	SET6	SET7	SET8	SET9	SET10	SET11
# var.	0	492	77	254	156	119	80	43	80	269	233	49
RMSE	2.866	**2.020**	2.396	2.193	2.243	2.264	2.308	2.421	2.261	2.046	2.055	2.306
Cor.	0.000	**0.711**	0.562	0.647	0.627	0.617	0.595	0.546	0.618	0.702	0.698	0.597
Adj. R^2	-0.000	0.659	0.552	0.617	0.608	0.602	0.585	0.540	0.608	0.675	**0.675**	0.590
U	0.881	**0.621**	0.736	0.674	0.690	0.697	0.711	0.742	0.695	0.629	0.632	0.710
U.bias	0.005	**0.001**	0.008	0.004	0.004	0.004	0.005	0.005	0.005	0.002	0.003	0.006
U.var.	0.997	**0.171**	0.284	0.203	0.215	0.224	0.241	0.305	0.224	0.176	0.180	0.248
U.cov.	0.000	**0.829**	0.709	0.795	0.783	0.773	0.756	0.691	0.772	0.823	0.819	0.747
RMSPE	15.850	19.186	15.845	18.174	17.128	18.227	16.141	**13.756**	18.616	19.444	20.262	16.182
CI.hit	0.515	**0.649**	0.596	0.604	0.596	0.584	0.564	0.612	0.594	0.640	0.637	0.583
Sign.	0.000	0.556	0.446	0.562	0.547	0.543	0.450	0.499	0.522	**0.567**	0.566	0.466
Neg2pos4	12.580	**8.415**	10.897	8.701	8.894	9.017	8.960	10.167	9.764	8.528	8.810	9.105
FsRMSE	11.561	6.831	7.763	6.800	7.048	6.898	7.186	8.565	8.414	6.946	7.025	6.747
Min. Dev.	-11.609	-10.589	-11.015	-10.430	-10.677	-10.730	-10.655	-10.913	-10.703	-10.564	-10.587	**-10.496**
Max. Dev.	17.870	**12.745**	14.829	13.628	13.645	13.808	13.986	15.162	14.255	12.883	12.908	14.210
Mean pos.	1.994	**1.376**	1.701	1.564	1.621	1.625	1.654	1.759	1.605	1.401	1.425	1.681
Mean neg.	-1.847	**-1.317**	-1.476	-1.414	-1.446	-1.481	-1.524	-1.385	-1.407	-1.337	-1.326	-1.484
Mean abs.	1.913	**1.342**	1.583	1.487	1.530	1.550	1.589	1.560	1.503	1.366	1.371	1.580
LPS	0.927	**0.565**	0.711	0.640	0.662	0.673	0.694	0.710	0.659	0.576	0.581	0.689
Adj. LPS	0.927	0.640	0.723	0.679	0.686	0.692	0.707	0.717	0.671	0.617	**0.616**	0.697
cEncomp.	0.000	0.923	0.076	0.043	-0.107	0.041	-0.005	0.098	-0.019	-0.028	0.081	-0.029
pEncomp.		**0.014**	0.251	0.387	0.433	0.452	0.564	0.213	0.267	0.495	0.354	0.477
Mean	**0.009**	-0.028	0.013	-0.068	-0.039	-0.048	-0.049	-0.070	-0.024	-0.049	-0.042	-0.068
Median	-3.268	**-2.615**	-3.135	-3.005	-3.090	-3.088	-3.156	-3.208	-3.030	-2.683	-2.650	-3.241
Sd	2.172	**1.587**	1.866	1.726	1.758	1.799	1.825	1.877	1.782	1.612	1.605	1.827
Skewness	-0.613	-0.128	-0.429	-0.457	-0.436	-0.473	-0.436	-0.715	-0.334	-0.180	**-0.095**	-0.565
Kurtosis	5.496	4.495	4.968	4.030	**3.888**	4.129	3.934	5.991	4.952	4.540	4.524	4.209

10 runs. Training set contains all studies. *SET0*: no variables. *SET1*: all variables. *SET2*: all significant variables (< 0.1) from SET1. *SET3-6*: EBA with criterion A-D. *SET7/8*: stepwise forward/backward (based on significance). *SET9/10*: stepwise backward/forward (based on AIC improvement). *SET11*: BMS (maximum of 49 variables at once). *# var.* is the number of variables used by each method. Bold cells are the best of each row. Rows below the line are statistics of the error term (excluding the smallest and largest 1% of the errors) and only the best (i.e., similarity towards $\mathcal{N}(0,\sigma^2)$) cells are marked.

Table 4.5: Classification ratings of the precision in predicting random studies

Method	SET0	SET1	SET2	SET3	SET4	SET5	SET6	SET7	SET8	SET9	SET10	SET11	N
# var.	0	492	77	254	156	119	80	43	80	269	233	49	
Sign	0.622	0.672	0.813	0.780	0.794	0.797	0.812	**0.844**	0.819	0.766	0.781	0.816	696
Pos. sign	0.270	0.442	0.386	0.490	0.481	0.487	0.464	0.455	0.477	0.518	**0.531**	0.486	178
Neg. sign	0.740	0.750	0.957	0.879	0.900	0.903	0.929	**0.975**	0.934	0.850	0.866	0.927	518
20% sign.	0.454	0.529	0.629	0.587	0.626	0.637	0.645	**0.654**	0.635	0.599	0.606	0.650	390
20% pos.	0.114	0.289	0.237	0.286	0.255	0.267	0.260	0.223	0.281	0.339	**0.349**	0.233	68
20% neg.	0.525	0.580	0.716	0.651	0.705	0.717	0.728	**0.747**	0.711	0.656	0.662	0.739	322
5% sign.	0.344	0.457	0.414	0.483	0.502	0.501	0.424	0.492	0.516	**0.521**	0.515	0.464	302
5% pos.	0.066	0.281	0.173	0.270	0.218	0.223	0.211	0.182	0.275	**0.333**	0.242	0.166	43
5% neg.	0.392	0.486	0.457	0.522	0.553	0.552	0.463	0.548	0.561	0.554	**0.562**	0.513	258

10 runs. Training sets contain random 90% of all studies (all other studies belong to the test sets). *SET0*: random guessing. *SET1*: all variables. *SET2*: all significant variables (< 0.1) from SET1. *SET3-6*: EBA with criterion A-D. *SET7/8*: stepwise forward/backward (based on significance). *SET9/10*: stepwise backward/forward (based on AIC improvement). *SET11*: BMS (maximum of 49 variables at once). *N* is the number of observations in each category of the test-set. *# var.* is the number of variables used by each method. Light/dark grey cells: 95%-CI does not contain the best average of the naive approaches (i.e., is better/worse than the best of SET0/1). Bold cells are the best of each row.

Table 4.6: Classification ratings of the precision in fitting random studies

Method	SET0	SET1	SET2	SET3	SET4	SET5	SET6	SET7	SET8	SET9	SET10	SET11	N
# var.	0	492	77	254	156	119	80	43	80	269	233	49	
Sign	0.622	0.861	0.823	0.835	0.833	0.832	0.833	0.848	0.845	**0.863**	0.862	0.832	696
Pos. sign	0.269	0.643	0.388	0.537	0.524	0.517	0.474	0.449	0.535	0.645	**0.647**	0.503	178
Neg. sign	0.740	0.934	0.969	0.937	0.937	0.938	0.954	**0.982**	0.949	0.936	0.934	0.942	518
20% sign.	0.454	0.676	0.658	0.667	0.669	0.671	0.661	**0.692**	0.646	0.683	0.665	0.658	390
20% pos.	0.114	0.322	0.232	0.305	0.289	0.295	0.266	0.217	0.267	**0.355**	0.330	0.228	68
20% neg.	0.525	0.751	0.751	0.744	0.750	0.751	0.745	**0.793**	0.727	0.752	0.737	0.750	322
5% sign.	0.343	0.556	0.446	0.562	0.547	0.543	0.450	0.499	0.522	**0.567**	0.566	0.466	302
5% pos.	0.066	**0.327**	0.163	0.294	0.239	0.292	0.197	0.165	0.286	0.305	0.315	0.153	43
5% neg.	0.392	0.596	0.497	0.609	0.601	0.589	0.495	0.560	0.566	**0.612**	0.610	0.520	258

10 runs. Training sets contain random 90% of all studies (all other studies belong to the test sets). *SET0*: random guessing. *SET1*: all variables. *SET2*: all significant variables (< 0.1) from SET1. *SET3-6*: EBA with criterion A-D. *SET7/8*: stepwise forward/backward (based on significance). *SET9/10*: stepwise backward/forward (based on AIC improvement). *SET11*: BMS (maximum of 49 variables at once). *N* is the number of observations in each category of the test-set. *# var.* is the number of variables used by each method. Light/dark grey cells: 95%-CI does not contain the best average of the naive approaches (i.e., is better/worse than the best of SET0/1). Bold cells are the best of each row.

do not belong to that category (given first in the second line) divided by the number of observations estimated to be in that category (given in the second line in parentheses). The number in parentheses in the first column is the actual number of observations in that category. The total miss rate is given in the third line and is calculated by the sum of the not correctly (the observation actually belongs to the category but not the estimate) and falsely classified (the estimate belongs to the category but not the actual observation) values divided by the total number of actual values belonging to that category (thus a 0 indicates a perfect classification while 2 is the worst case when every observation is incorrectly classified[7].

Looking at table 4.7 the stepwise regressions (SET7 and SET8) appear to perform best while it is appropriate to point out that all estimators do much better than the naive approaches. Not surprisingly, the number of positive outcomes is generally underestimated and the percentage of wrong classifications is better for negative (normalized) t-values than for positive values. It is interesting to note that the number of positive results is underestimated by the stepwise regressions based on the significance levels while it is overestimated by the stepAIC algorithm. For SET7 and SET8 the total miss rates are all below one. Although the EBA results are not as good, they perform fairly well in predicting the sizes of the categories.

Again, the picture changes when all studies are used to construct the estimators (table 4.8). SET1 performs better but is the best in just one category. SET7 and SET8 fall back behind SET9 and SET10. All in all, the total miss rates have been reduced by a large margin. Surprisingly, the estimated number of observations in each category has, overall, become worse.

4.2.3 And the Winner...

There is no more common error than to assume that, because prolonged and accurate mathematical calculations have been made, the application of the result to some fact of nature is absolutely certain.

Alfred N. Whitehead, Alfred North Whitehead: An Anthology, 1953

...depends on the aim of the researcher. Shall the estimator fit (ex post prediction) the existing data as well as possible? Or should the estimator predict (ex ante) unknown data? Is a good fit more or less important than a general classification?

One general conclusion seems to be that selecting fewer variables is better for predicting but worse for fitting the data. Bayesian Model Averaging was restricted to 50 variables and stands out by its high variability: in some runs its predictive/fitting performance is very good and in some cases it is extraordinarily bad. This may come from utilizing too much detailed information from some studies which are rather specialized and not suited to be used for other studies because

[7]This is the case when every value which is estimated to belong to that category actually does not, and every observation which does not belong to that category is estimated to be part of it.

Table 4.7: Classification ratings of the errors in predicting random studies

Method	SET0	SET1	SET2	SET3	SET4	SET5	SET6	SET7	SET8	SET9	SET10	SET11
# var.	0	492	77	254	156	119	80	43	80	269	233	49
Sign	0.378	0.328	0.187	0.220	0.206	0.203	0.188	**0.156**	0.181	0.234	0.219	0.184
# false (696)	263 (696)	228 (696)	130 (696)	153 (696)	144 (696)	141 (696)	131 (696)	**108** (696)	126 (696)	163 (696)	152 (696)	128 (696)
Total miss rate	0.756	0.655	0.373	0.440	0.413	0.407	0.376	**0.311**	0.363	0.468	0.437	0.369
Pos. sign	0.741	0.632	0.223	0.420	0.378	0.361	0.298	**0.133**	0.289	0.461	0.425	0.304
# false (178)	131 (177)	154 (243)	30 (136)	75 (179)	67 (177)	64 (178)	34 (114)	**10** (76)	42 (144)	101 (219)	85 (201)	37 (121)
Total miss rate	1.466	1.423	0.785	0.933	0.895	0.875	0.727	**0.602**	0.756	1.050	0.949	0.722
Neg. sign	0.249	0.200	0.176	0.164	0.163	0.162	0.163	0.158	0.158	0.161	**0.154**	0.157
# false (518)	129 (519)	90 (452)	98 (559)	84 (516)	85 (518)	84 (517)	95 (582)	98 (619)	87 (552)	77 (476)	**76** (495)	90 (574)
Total miss rate	0.510	0.425	0.233	0.284	0.263	0.258	0.254	**0.214**	0.235	0.298	0.281	0.247
Sig20	0.543	0.512	0.332	0.388	0.348	0.343	0.329	0.317	**0.301**	0.373	0.356	0.318
# false (390)	239 (440)	237 (463)	124 (375)	144 (371)	125 (359)	122 (356)	125 (379)	125 (394)	**105** (348)	143 (382)	137 (385)	124 (389)
Total miss rate	1.158	1.080	0.690	0.782	0.695	0.676	0.675	0.667	**0.634**	0.766	0.745	0.667
Pos. sig20	0.885	0.813	0.437	0.617	0.609	0.468	0.400	0.450	**0.393**	0.623	0.582	0.549
# false (68)	66 (74)	117 (144)	25 (57)	44 (71)	36 (59)	25 (54)	**10** (24)	**10** (23)	17 (44)	61 (97)	43 (74)	14 (25)
Total miss rate	1.853	2.437	1.131	1.356	1.270	1.102	**0.881**	0.930	0.970	1.553	1.281	0.972
Neg. sig20	0.476	0.417	0.323	0.347	0.318	0.325	0.321	0.305	**0.288**	0.323	0.313	0.303
# false (322)	174 (365)	133 (319)	103 (318)	104 (300)	95 (300)	98 (303)	114 (356)	113 (371)	**88** (304)	92 (285)	97 (311)	110 (363)
Total miss rate	1.015	0.833	0.603	0.672	0.591	0.589	0.626	0.604	**0.561**	0.630	0.641	0.603
Sig95	0.610	0.579	0.396	0.443	0.380	0.380	0.383	0.350	**0.348**	0.408	0.391	0.362
# false (302)	194 (319)	217 (375)	78 (198)	113 (255)	91 (238)	88 (232)	80 (210)	76 (216)	**75** (217)	103 (253)	103 (264)	80 (221)
Total miss rate	1.299	1.262	0.846	0.892	0.798	0.792	0.842	0.759	**0.734**	0.821	0.826	0.801
Pos. sig95	0.923	0.841	0.510	0.634	0.726	0.567	0.518	0.484	**0.419**	0.614	0.616	0.634
# false (43)	39 (43)	95 (113)	15 (30)	32 (51)	22 (30)	15 (26)	9 (17)	**6** (13)	9 (21)	34 (55)	25 (40)	9 (14)
Total miss rate	1.850	2.923	1.180	1.484	1.293	1.120	0.993	0.960	**0.930**	1.454	1.333	1.038
Neg. sig95	0.564	0.501	0.390	0.412	0.355	0.370	0.375	0.341	**0.341**	0.373	0.360	0.352
# false (258)	156 (276)	131 (262)	**66** (168)	84 (204)	74 (208)	76 (206)	72 (193)	69 (203)	67 (196)	74 (198)	80 (224)	73 (207)
Total miss rate	1.211	1.022	0.797	0.803	0.733	0.745	0.817	0.721	**0.698**	0.732	0.749	0.769

10 runs. Training sets contain random 90% of all studies (all other studies belong to the test sets). *SET0*: random guessing. *SET1*: all variables. *SET2*: all significant variables (< 0.1) from SET1. *SET3-6*: EBA with criterion A-D. *SET7/8*: stepwise forward/backward (based on significance). *SET9/10*: stepwise backward/forward (based on AIC improvement). *SET11*: BMS (maximum of 49 variables at once). *# var.* is the number of variables used by each method. Light/dark grey cells (first row of each category): 95%-CI does not contain the best average of the naive approaches (i.e., is better/worse than the best of SET0/1). Bold cells are the best of each row. # false: the number of falsely estimated data in the corresponding category; the number in parentheses is the estimated number of observations in that category (the number in the first column is the actual number). The total miss rate is the sum of errors of both kinds divided by the actual number of observations in the corresponding category.

Table 4.8: Classification ratings of the errors in fitting random studies

Method # var.	SET0 0	SET1 492	SET2 77	SET3 254	SET4 156	SET5 119	SET6 80	SET7 43	SET8 80	SET9 269	SET10 233	SET11 49
Sign	0.378	0.139	0.177	0.165	0.167	0.168	0.167	0.152	0.155	**0.137**	0.138	0.168
# false (696)	263 (696)	96 (696)	123 (696)	115 (696)	116 (696)	117 (696)	116 (696)	105 (696)	108 (696)	**95**	96 (696)	117 (696)
Total miss rate	0.756	0.277	0.355	0.330	0.334	0.337	0.334	0.303	0.309	**0.274**	0.275	0.336
Pos. sign	0.741	0.232	0.185	0.256	0.260	0.263	0.214	**0.098**	0.228	0.227	0.230	0.258
# false (178)	131 (176)	34 (145)	20 (111)	33 (128)	35 (134)	34 (128)	23 (108)	**7** (71)	27 (118)	33 (145)	33 (145)	31 (129)
Total miss rate	1.465	0.546	0.727	0.648	0.671	0.672	0.656	0.590	0.617	**0.540**	0.541	0.671
Neg. sign	0.249	0.114	0.174	0.144	0.147	0.148	0.157	0.159	0.140	0.113	**0.112**	0.150
# false (518)	129 (519)	63 (550)	102 (585)	81 (567)	82 (562)	84 (567)	92 (587)	99 (625)	81 (577)	**62** (550)	**62** (550)	86 (575)
Total miss rate	0.510	0.187	0.228	0.220	0.222	0.224	0.223	0.209	0.208	**0.184**	0.185	0.224
Sig20	0.544	0.265	0.310	0.280	0.286	0.309	0.305	0.299	0.274	**0.261**	0.261	0.302
# false (390)	239 (440)	95 (358)	106 (342)	94 (337)	96 (334)	105 (338)	119 (391)	121 (404)	**86** (314)	93 (355)	95 (364)	116 (383)
Total miss rate	1.159	0.567	0.614	0.574	0.576	0.598	0.645	0.618	0.575	**0.555**	0.578	0.638
Pos. sig20	0.884	0.411	**0.327**	0.376	0.344	0.380	0.375	0.389	0.330	0.399	0.406	0.452
# false (68)	65 (74)	16 (38)	**5** (16)	11 (30)	9 (27)	11 (29)	8 (21)	8 (20)	7 (22)	16 (40)	17 (43)	9 (20)
Total miss rate	1.842	0.910	0.844	0.859	0.849	0.869	0.852	0.894	**0.840**	0.878	0.924	0.903
Neg. sig20	0.476	0.244	0.309	0.269	0.280	0.299	0.304	0.290	0.270	0.240	**0.240**	0.293
# false (322)	174 (366)	78 (319)	101 (326)	82 (307)	86 (307)	93 (309)	112 (370)	112 (385)	79 (292)	**76** (315)	77 (321)	107 (363)
Total miss rate	1.017	0.491	0.562	0.512	0.516	0.537	0.603	0.554	0.517	**0.483**	0.502	0.581
Sig95	0.610	0.287	0.323	0.300	0.306	0.327	0.358	0.339	0.334	**0.280**	0.291	0.347
# false (302)	194 (319)	64 (22)	**57** (176)	64 (212)	68 (222)	74 (225)	74 (206)	75 (221)	65 (194)	59 (211)	65 (224)	76 (218)
Total miss rate	1.300	0.656	0.742	0.649	0.678	0.701	0.795	0.749	0.692	**0.629**	0.650	0.784
Pos. sig95	0.922	**0.311**	0.492	0.351	0.442	0.433	0.429	0.465	0.404	0.318	0.338	0.605
# false (43)	39 (42)	6 (19)	**4** (9)	6 (17)	7 (15)	7 (16)	6 (14)	5 (11)	5 (13)	5 (16)	6 (18)	7 (12)
Total miss rate	1.832	0.814	0.936	0.842	0.918	0.872	0.944	0.950	0.838	**0.813**	0.826	1.011
Neg. sig95	0.565	0.279	0.314	0.294	0.295	0.324	0.353	0.333	0.334	**0.273**	0.286	0.341
# false (258)	156 (277)	57 (57)	**53** (167)	57 (196)	61 (207)	68 (209)	68 (192)	70 (211)	60 (181)	**53** (195)	59 (206)	70 (206)
Total miss rate	1.215	0.624	0.707	0.613	0.635	0.674	0.768	0.712	0.667	**0.594**	0.618	0.753

10 runs. Training sets contain random 90% of all studies (all other studies belong to the test sets). *SET0*: random guessing. *SET1*: all variables. *SET2*: all significant variables (< 0.1) from SET1. *SET3-6*: EBA with criterion A-D. *SET7/8*: stepwise forward/backward (based on significance). *SET9/10*: stepwise backward/forward (based on AIC improvement). *SET11*: BMS (maximum of 49 variables at once). *# var.* is the number of variables used by each method. Light/dark grey cells (first row of each category): 95%-CI does not contain the best average of the naive approaches (i.e., is better/worse than the best of SET0/1). Bold cells are the best of each row. # false: the number of falsely estimated data in the corresponding category; the number in parentheses is the estimated number of observations in that category (the number in the first column is the actual number). The total miss rate is the sum of errors of both kinds divided by the actual number of observations in the corresponding category.

Bayesian Model Selection, which is inherently similar to BMA, performs quite good in predicting the data and only slightly worse in fitting them. In our scenario, BMA does not perform as well as in Fernández et al. (2001b), who compare the naive estimator, one EBA version and BMA in the case of a moderately sized set of variables in the context of country growth. Their conclusion, that BMA is superior to EBA, cannot be unambiguously confirmed.

EBA performs mediocre in most cases except for predicting the sizes of various general classes. However, we have to keep in mind that EBA only selects seemingly important and stable variables but does not - because we restricted the number of variables to be included in every regression - evaluate their conjoint influences as the other approaches do.

All in all, the stepwise approaches perform best, while those based on the significance levels of each variable (which lead to fewer variables) are good in predicting unknown results. Those based on the model improvement are more suited to fit the data. This applies both to the general fit of the estimators as well as to their classification performance.

Since the out of sample prediction is commonly the preferred way to judge a forecasting method, we come to the following conclusions. In the case of our data set, we prefer stepwise regressions to select the variables for further inspection. In general, when dealing with data sets containing numerous variables in an exploratory context, it seems advisable to employ methods which resort to a small set of really important variables. Including too many variables seems to dilute findings. It may not be a good idea to rely solely on "expert opinions" in selecting variables, because several important influences could be missed since they do not belong to any underlying theory.

4.3 Further Results from the Best Models

Mathematics is not a careful march down a well-cleared highway, but a journey into a strange wilderness, where the explorers often get lost. Rigour should be a signal to the historian that the maps have been made, and the real explorers have gone elsewhere.

Anglin (1992)

Since we have chosen the stepwise estimators to be the most useful, we may delve deeper into their results than we have already done in subsection 3.6.5. Besides the results already given there, several other implications can be derived from table 3.52. In both regressions the significant coefficients of the number of the ending page are larger in absolute value than of the starting page of a study (-0.006112 and -0.006064 versus 0.005689 and 0.005788; the differences of the absolute values are also significant). This implies that studies which cover more pages report results which are more negative than shorter studies; studies with a larger starting page also yield more negative values. However, the page numbering is not always consistent between journals because some issues of one volume are consecutively numbered while others are not.

The number of measuring points (in time) also shows a significant positive statistical linkage to the results. However, because the coefficients (0.0005378 and 0.0005834) are quite small and the median number of measuring points is just three (since surveys, experiments and cross-sections usually have only one measuring point in time), this variable seems to be catching just some anomalies in the data. The time a study covers (measured in month; coefficient is -0.00074 in both regressions) is significantly related with smaller (normalized) t-values. Although this implicates that analyzing a longer period in time is associated with more significant results in favor of the deterrence hypothesis, we have to keep in mind that this variable is, to some extent, a substitution for studies using time-series; and the latter variable is not included in any of the four stepwise regressions.

It is somewhat surprising that the population sizes are included and very significant in every regression although the number of observations (as the square root and diversified by its sign) is also included as well. The inclusion of significance-tests is negatively related on the study-level and positively on the estimate-level. Since the absolute coefficients are very similar this indicates that estimates without a value of significance in a study which provides such tests in general are more negative. On the other hand, this may be just be an artifact of our transformation scheme[8].

Compared to books and conference papers, estimates published in journals yield results which are less in compliance with the deterrence hypothesis. The opposite is observed for working and discussion papers and other types of publication. Economists and psychologists are more likely to produce results which are more in line with the deterrence hypothesis while the contrary is observed for sociologists and criminologists (as can be seen by the discipline of the author or the journal). However, some authors are included with deviating signs[9]. Among these, the most prominent economists seem to be Levitt, Ehrlich, Cloninger and Witte. Although publishing in an economic journal is accounted for, Levitt, Cloninger and Witte bear a positive sign which could imply that “their” results are, after controlling for many other influences, less in favor of the deterrence hypothesis. This is especially interesting for Steven Levitt because he is known for his innovative ideas of finding evidence of deterrent effects. For criminologists and sociologists the picture is less clear.

Amongst the studied countries two conclusions seem to be possible[10]. Studies with German, Swedish, Australian and Dutch data are associated with more negative (normalized) t-values while the opposite is observed for studies using data from the USA, UK, Canada and other countries. Especially Canada stands out by being included in every regression with large and very significant positive coefficients. In regard to studies using German data we have to remark that the coefficient of using the PKS as a public data base is positively signed and is a very large - in contrast to the coefficient of studying German data, which is negative and is somewhat lower in absolute value.

[8]Such a t-value has to be a transformed test statistic.

[9]We have to keep in mind that the reference category is, depending on the actual regression, quite different and relates to several authors.

[10]All countries are listed in table 3.52; hence, the reference category depends on the actual regression.

In regard to the theory of a U-shaped crime distribution (Howsen and Jarrell, 1987), it is interesting to note that deterrence is more readily found to be effective for very small and large locations. Thus, not only does the crime density increases in very densely and sparsely populated areas but also does the significance of deterrent effects.

While the number of bivariate results is positively related to the (normalized) t-values, no similar relationship is observed for the number of multivariate results. Although only a few observations exist, studying social fringe groups is accompanied by large positive coefficients. Several technical variables in regard to representativeness, reliability and pretests are included as well, but their proper interpretation seems to be difficult.

Studies which do not test predefined deterrence hypotheses are associated with more positive results. Moreover, as can be seen in the other models, results are more negative when the studied deterrence variable is the focus of the study. Whether or not this can be interpreted in the context of publication bias remains unclear but plausible. It also seems to be that smaller (normalized) t-values can be observed for subsets which consist of males, youths, adults or people from rural areas; however, the number of such observations in our meta-data base is very small. As before, (normalized) t-values in regard to the deterrent effect of the death penalty are much larger than all others. They are especially large when the deterrent effect is measured by the existence of the death penalty.

Among the deterrence measures several general issues can be observed: there are only very few variables with negative coefficients (conviction rate, convicted per crime, arrest rate and incarcerated per crime); all other variables are positively signed, even the other incarceration measures. Especially, all measures in regard to the severity of punishment have a positive sign. The same picture emerges for surveys: the probabilities of detection and punishment are negatively signed while the algorithms have included only one measurement of the severity of punishment which bears a positive sign (severity of punishment by others). It is noteworthy that the coefficients for the probabilities are much larger (in absolute value) when detection or punishment relates to friends or the family (in contrast to the police, justice or employment). This might be interpreted as evidence that social capital may be even more important than human capital. Last but not least, only the probability of detection (with a negative sign) is included in the case of experiments.

Among the endogenous variables which measure the number of offenses only a few are significant and are included: violating prescriptive limits and self reported delinquency (lifelong). Since the reference category consists of several other measures, interpretation is somewhat difficult. However, it is easier in the case of the general crime category: only for misdemeanors the deterrent effects are more significant than for crimes. Among the studied crimes there are also some interesting observations. We have already mentioned that dealing with hard drugs bears the opposite sign than dealing with soft drugs. The same can be observed for assault and negligent assault which might implicate that negligent offenses are less affected by deterrence measures. Since each estimate is often based on multiple offenses, a more detailed interpretation is difficult and the question why so many of these variables are positively signed remains unanswered.

The situation is similar in the case of the implemented covariates. However, it is interesting that the absolute coefficients of the profession, social class, drug usage, morality, random effects, GDP and risk propensity are, compared to all others, quite large. Especially the latter has a coefficient twice as large as the second largest. It might be a coincidence but most of the economic covariates (GDP, labor force and consumption) are negatively signed while most social and personal covariates have a positive sign.

The dummy which indicates that an estimate is not corrected for simultaneity is negatively signed and highly significant. Compared to those estimates which do use methodological means to correct for simultaneity (the reference category), relying only on additional variables - usually lagged variables - is also associated with smaller (normalized) t-values. This could mean that results relying on models which do not consider simultaneity problems overestimate deterrent effects. This bias is reduced partially when feedback effects are taken care of with additional variables. However, this interpretation is put into perspective when we look at the highly significant and negative coefficients of the dummy which indicates the usage of 2SLS because methodological correction for simultaneity and 2SLS are highly correlated. GLS and ARIMA models are also highly significant and bear a negative sign.

As is the case in all other models, the square roots of the number of observations each estimate is based on (diversified by their sign) are both highly significant. This underlines that the statistical relationship between the values of significance and the number of observations is not negligible and should be considered in a meta analysis. This should be done in regard to publication bias as well to explain more variance.

5 Conclusion

An education isn't how much you have committed to memory, or even how much you know. It's being able to differentiate between what you do know and what you don't. It's knowing where to go to find out what you need to know; and it's knowing how to use the information you get.

William Feather

The most merciful thing in the world, I think, is the inability of the human mind to correlate all its contents. We live on a placid island of ignorance in the midst of black seas of infinity, and it was not meant that we would voyage far. The sciences, each straining in its own direction, have hitherto harmed us little; but some day the piecing together of dissociated knowledge will open up such terrifying vistas of reality, and of our frightful position therein, that we shall either go mad from revelation or flee from the deadly light into the peace and safety of a new dark age.

Howard P. Lovecraft, The Call of Cthulhu, 1926

The main result most certainly is that the outcome of an empirical study of the deterrent effect of punishment is not independent of its design. It is not surprising that the kind of offense or the studied population play an important role. However, it is not self-evident that personal characteristics of the authors, the implemented techniques, estimation properties or the design of a study in general also have an important impact on the results. But let us begin in a chronological order.

This study emerged from a project financed by the DFG and conducted by two teams from different universities. After multiple stages of filtering, we identified - from a pool of more than nine thousand studies found by a preceding search in numerous data bases - 840 empirical studies which contain relevant results in regard to deterrence. A large set of information about each study - including characteristics of the author and publisher, the design of the study and its techniques - was entered into a shared data base by both teams. Due to time constraints each team could only process 350 studies; a random set of 140 studies had to remain unattended. The results of each study were - if not already given explicitly - recoded or transformed into normalized[1] t-values. Subsequently, we analyzed these (normalized) t-values on their own and their relationship with all other recorded information in a meta regression analysis.

While interpreting the derived results, we have to keep in mind that the estimates of these studies are somewhat distorted. Although the evidence of publication bias is not sufficient to

[1]The dependence on the degrees of freedom was removed and their sign made consistent.

correct the data, the apparent sparse distribution in some regions of the (normalized) t-values indicates that some bias is present. It should be less of a problem if the models were respecified until their results look good because this is taken account of in our meta analysis. Nevertheless, a publication bias may have other sources (intentional or unintentional). Therefore, the average (normalized) t-value of -1.4 (median -1.37) should be interpreted carefully. It might be slightly biased downwards (i.e., it is too small) but, on the other hand, it is averaged over all studies and results. For example, as shown in section 3.5, studies conducted in the United Kingdom or Germany yield results which are more in line with the deterrence theory than those studies employing data from Canada (table 3.31). Another example is the type of offense (table 3.43): while results for tax evasion, drunk driving and fraud - and property crime in general - are more compatible with the deterrence hypothesis, those for homicide or assault are not. Results for the probability of punishment are also more in favor of the deterrence hypothesis than those based on the severity of punishment. It also seems important to note that studies focussing on deterrence report results which are slightly but significantly more supportive of the deterrence hypothesis. However, in every analysis those variables associated with the sample size are highly significant. This also indicates, as explained in section 3.4, that some effect exists. All in all, summing over all studies, the null-hypothesis of no deterrent effect should be discarded rather than accepted. This view is also consistent with the overall opinion of the authors.

In a bivariate analysis we focus on comparisons of several selected variables we assume to be important. Among other things, we find that when the authors or the journal belong to the field of psychology or economics their results are significantly more in line with the deterrence hypothesis than those from sociological or criminological authors or journals. As shown in table 3.36, the results also vary largely between the most frequent authors. While there is no trend in the year of publication, there is weak evidence that studies using newer data yield results which are less in favor of the deterrence hypothesis. Moreover, studies using experiments do agree more with the deterrence theory than those using time series, panel or cross sectional data. The same applies when studies using no public data base are compared to those using the UCR.

When dealing with certain properties of the estimators there are also some distinctive features: using the crime rate as the endogenous variable yields results which are less in favor of the deterrence hypothesis than most other definitions like the reported delinquency, accidents, etc. When analyzing crime data the same is observed for exogenous variables: comparing police, inspection variables or sentence lengths to arrest rates or convictions per crime, the latter are more in line with the deterrence theory. In regard to surveys and experiments the same is observed for variables measuring the severity of punishment and the probability of detection. The studied offense also seem to be important: while analyzing speeding, tax evasion, drunk driving, larceny, fraud or environmental offenses, the reported estimates are supporting the deterrence hypothesis much more than in the case of homicides, drug dealing, assault or the general crime rate. Furthermore, there is evidence that the employed techniques play an important role. Using Pearson correlations, path analysis or other multivariate (those not listed in table 3.48) methods yield results which are less supportive of the deterrence theory. Tobit analysis or GLS shows the contrary. For some techniques (e.g., VAR methods) it is important whether or not simultaneity is considered; however the effect is not consistent for all methods.

Although such bivariate comparisons have their merit one should not rely on them exclusively. Some of the previously described effects vanish or are even reverted (e.g., the effect of some authors like Piquero), when several variables are considered simultaneously in a multivariate setting. We employ several methods to identify important variables and their influence: simple OLS, EBA, stepwise regressions and bayesian estimators. To judge the quality of these methods their performance in fitting and predicting the data, as well as their precision therein, is compared. We conclude that the stepwise regression estimators are of superior quality in our case. However, there are some issues all models agree upon: when studies provide tests of significance, deterrence is their focus, are written by Ehrlich, deal with the possession of hard drugs, use a dummy as the endogenous variable or do not correct for simultaneity, than their results are significantly more in favor of the deterrence hypothesis. An opposite effect can be observed when Canadian people are studied, the representativeness of the data is not checked, the death penalty is studied or when poverty or welfare or urbanity or population (-growth) is used as a covariate.

When we resort to stepwise regression only we see that, additionally to those results just mentioned, the following properties imply more negative (normalized) t-values: being an economic study, not being published in a journal, the author is Rasmussen or Hakim or Virén, large cities or (small) villages are studied, the probability of detection and punishment (by an institution or friends or family) is used in surveys, the probability of detection is used in experiments, the exogenous variable refers to a whole year, misdemeanors or "other" offenses are studied, age or social class or drug usage or spatial fixed effects or GDP or consumption is used as a covariate, bivariate regression or point biserial correlation (however, the sign on Pearson correlation is indecisive) or 2SLS or GLS or MANOVA or an "other" multivariate method is employed.

By contrast, larger (normalized) t-values are observed when the author is from Sweden or Switzerland or the Netherlands, the study is published in a journal, the author is Levitt, the study is of criminological or sociological nature (however, there are differences within such journals), the exogenous variable is the clearance rate or a sentence length or an incarceration rate or a police variable, U.S. data is used, information about the data base is not missing, state-based data is used, the exogenous variable is used in differences, crime is measured with recidivism, convicted to prison sentences or accidents, the endo- and exogenous variable do not relate to the same offense, possession of soft drugs or environmental offenses are studied, the offense has characteristics of violent and property crime, marital status or morality or random effects or alcohol consumption or risk propensity is implemented as a covariate, or binomial tests or COX regressions are used.

To sum up, there are several properties which significantly add to the determination of a study's results. This implies that the design of an empirical study and its authors determine - at least to some extent - the published results. Furthermore, the large variety of results found in the literature about the deterrent effect may partially be explained by the variety of implemented methods, study designs, employed sources and other properties. It would not be surprising when this also applies to other interesting topics. However, although the size of empirical coverage in deterrence research may not be unique, it is certainly an exception.

Even though this thesis has become very long there still remain aspects which deserve more attention. The applied techniques have analytical deficiencies as they only cover linear relationships, all eventually rely on OLS and their applicability is not perfect (as far as the usual assumptions about the error terms are concerned). There are also several technical problems of which the computing time and its scaling in the number of observations is most prominent. All EBA methods, as well as the bayesian models, suffer from the restriction on a certain number of variables. Beside other regression methods like MMA or an optimization of the bayesian models (using other priors or other model weights) the usage of non-linear methods could prove useful. Among such methods, decision trees and RSDA have been previously tested with smaller test sets.

It is inherent to the data we use that different results may be derived therefrom. While some results are derived from every method, some are not. Using the prediction quality and coverage is just one way to judge the reliability of the estimators. We do not explicitly rule out that there may be better ways to do so but we are confident that our *tournament* is a useable benchmark.

We are sure that further insights may be reached when the analysis is focussed on certain subsets instead of using all data simultaneously. The subsets described in section 2.2 should be promising candidates; especially the death penalty, drunk driving, tax evasion, environmental offenses and Index I crimes. The established data base includes enough literature covering these subjects to make such a specialized meta analyses viable. Furthermore, we have concentrated on the direction of the influential variables and their significance but did not discuss the seize of their coefficients and how these should be judged.

So, who benefits most from this thesis? What guidelines can be drawn for public policy? The derived results should significantly improve the understanding of the existing studies which estimate the deterrent effect of the probability or severity of punishment. There have been many ongoing discussions about the reliability of many results, the "correct" handling of the data and their interpretation. A major contribution of this study is the provision of a guidance to properly associate the discussion with the context. At least some discrepancies found in the literature can be traced back to properties of the author, the design of the study and its estimators. Furthermore, there is evidence that - considering all studies simultaneously - the existence of a deterrent effect is much more likely than its absence. Although this may not stop the discussion about the existence of a deterrent effect of punishment, this work is certainly a good and important argument.

Another aspect is the technical conception of this study which may be used as a guideline for further meta regression analyses. This applies to the acquisition of the literature, the meta regression methods, as well as their comparison and interpretation. The whole concept - or parts of it - are applicable to many other fields in which a meta analysis is useful.

A Appendices

One of the most fortunate situations a scientist can encounter is to enter a field in its infancy.

Bernhard Schölkopf and Alexander J. Smola, Learning with Kernels, 2001

Some statistics and tables are not really necessary to understand the full scope of the study. Some parts, like A.1 and sections A.2, are just additional information and statistics about some sections and further information which go beyond the central theme of this study.

On the other hand, section B.1 shows the basic statistics of all variables. Although these lists are quite long, providing these seems essential for such a study. Finally, section B.2 gives a list of all relevant and included studies accompanied by the opinion of their authors, their date of publication and who entered the information into the data base.

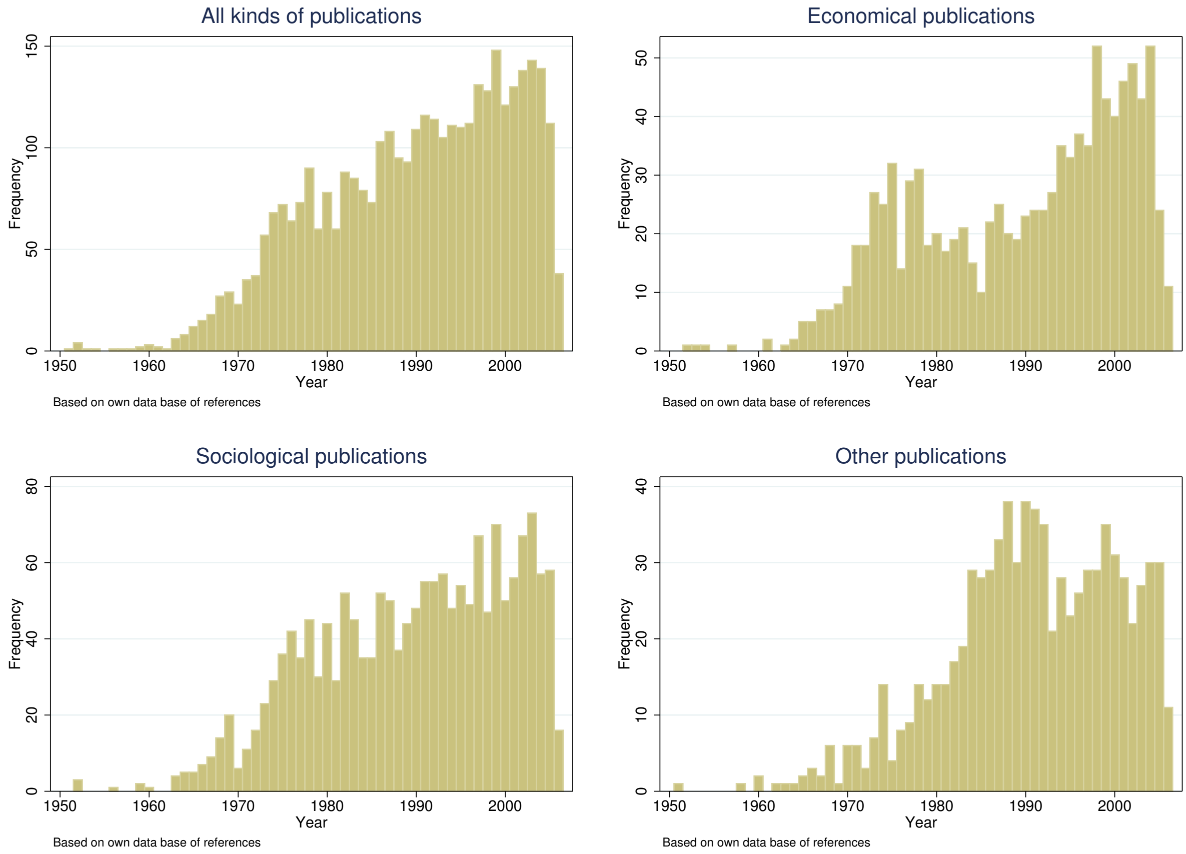
A.1 Other Statistics

A.1.1 Statistics About the References

As in many other fields of science the number of publications has steadily increased over time, as seen in figure A.1 (based on the 3598 references after the first stage of the search process; refer to subsection 3.1.1). In fact, this is not true for all disciplines. Whereas sociological studies steadily increased over time, economic publications decreased by a large amount in the eighties and recovered in the nineties, while "other" studies had a short temporary increase in the late eighties. The drop in the field of economics could be explainable by a receding interest in explaining the deterrence model after the long debate about Ehrlich's studies which started in the mid seventies and ebbed in the eighties. Then, it became popular to include deterrence measures as covariates in other studies; e.g., to understand the crime-unemployment relationship (Entorf and Winker, 2002). It is also interesting to note that the non-economical and non-sociological literature started to study deterrence after a lag of ten years (instead of the mid seventies, the main surge began in the mid eighties).

It is not surprising that, the number of pages per study has increased over time, as can be seen in figure A.2, and is significantly higher for sociological studies. Regressing the number of pages on the year and dummies for economic and sociological studies (taking all other disciplines as the reference) for articles results in (t-vales in brackets): $\text{pages} = 0.21(8.66)\text{year} +$

Figure A.1: Number of publications in the course of time (1950-2006)



$0.45(0.56)\text{economic} + 4.4(6.27)\text{sociologic} - 401(-8.30)$ with $N = 2505$, adj. $R^2 = 0.05$. The picture remains largely the same if we look only at relevant studies in our data base: pages $= 0.27(7.69)\text{year} + 2.31(2.29)\text{economic} + 5.09(5.46)\text{sociologic} - 520(-7.46)$ with $N = 697$ and adj. $R^2 = 0.10$.

In this context, it is also worth mentioning that there seems to be, when looking at the median (normalized) t-value, an interesting relationship between the number of pages and the associated approval of the deterrence theory. As the number of pages of a study increases, the median (normalized) t-value increases as well (i.e., the deterrence theory is less supported). This is illustrated by table A.1 (the categories are chosen in such a way that each contains approximately the same amount of estimates). However, these differences are not significant and can be explained by the finding that sociological studies are less supportive of the deterrence hypothesis and tend to consist of more pages. When considering other factors this relationship is even reversed (see section 4.3).

A.1.2 Increasing Efficiency of the Literature Data Bases Over Time

In the nineties many (and now most) publishers are present in the internet and their publications are administered electronically. The same applies for authors who publish their working papers

Figure A.2: Number of pages per study in the course of time (1950-2006)

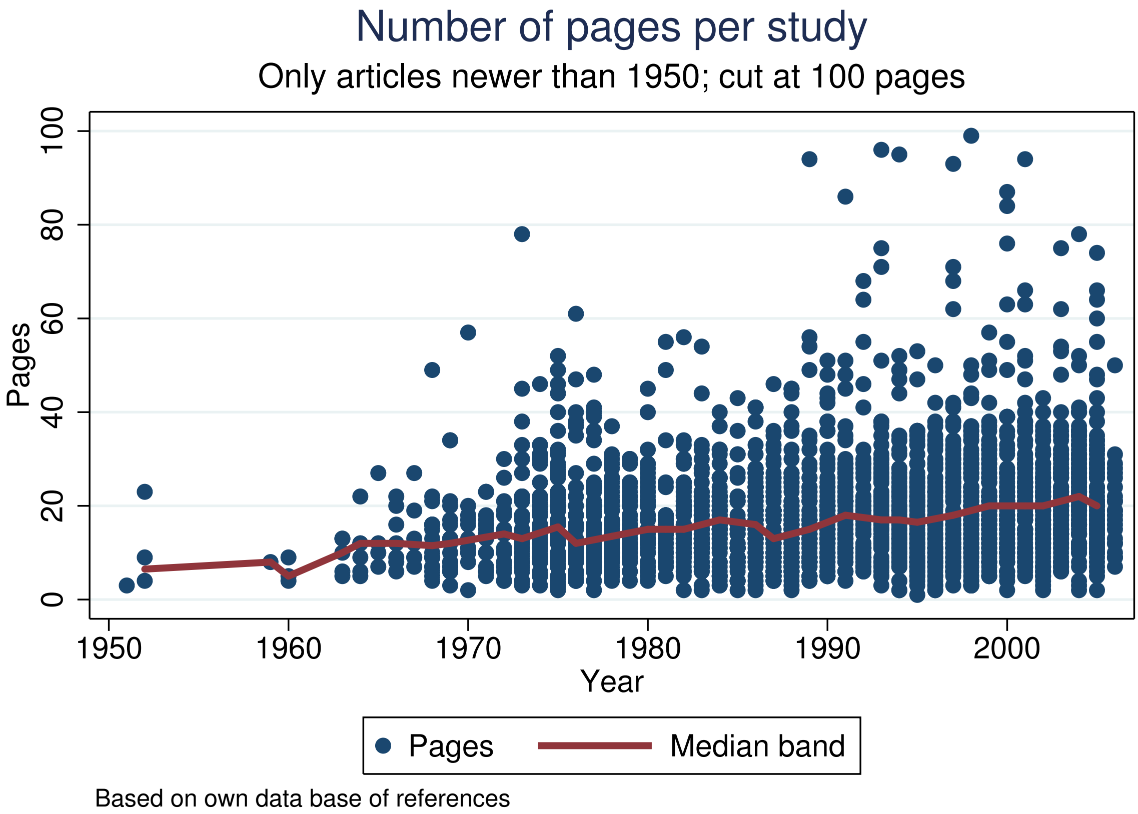


Table A.1: Differences by the number of pages

Pages	mean	median	%	#e	#s
1-9	−1.40	−1.73	48.03	1077	109
10-13	−1.33	−1.45	41.01	1160	119
<u>19-27</u>	−1.37	−1.35	40.53	1520	155
Overall mean	−1.40	−1.37	41.66	6530	663
14-18	−1.41	−1.33	41.31	1394	139
28-439	−1.51	−1.11	38.81	1379	141

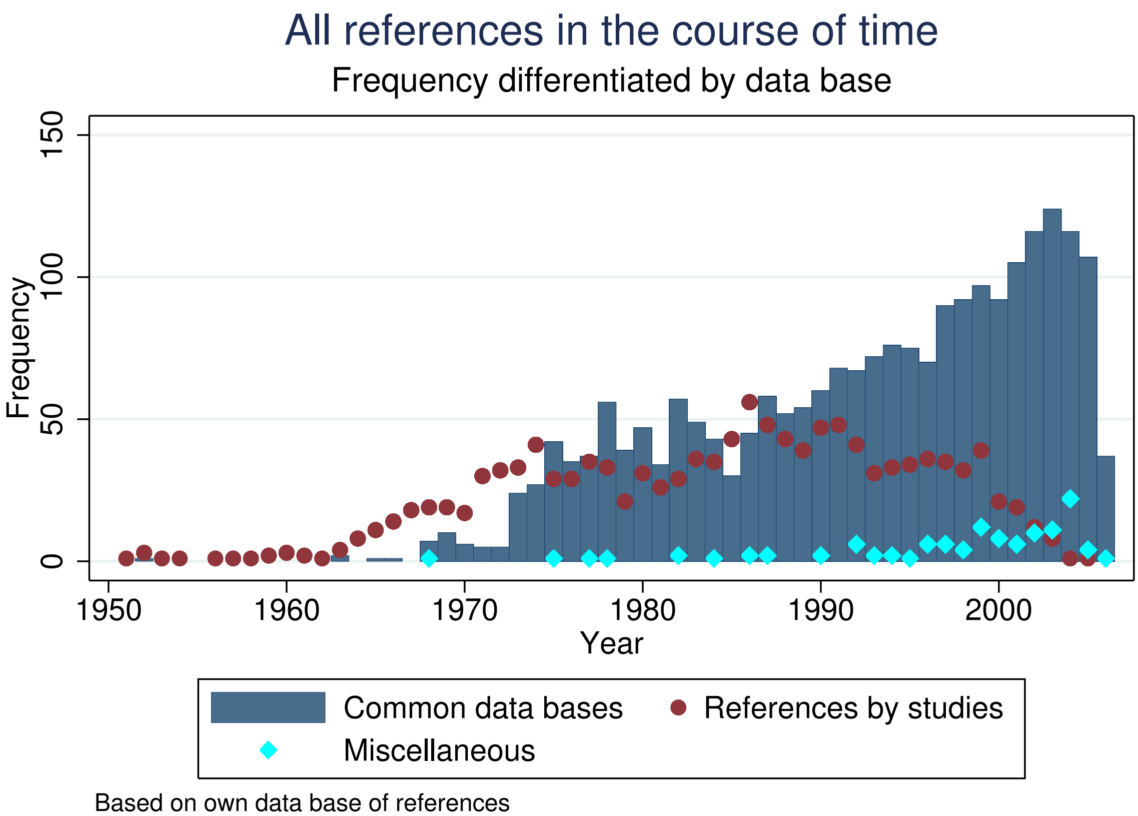
Mean and median correspond to the (normalized) t-values of the particular group. % is the percentage of estimates which are consistent with the deterrence hypothesis and significant at a 5% level in a two sided test. *#e* is the weighted number of all valid estimates. *#s* is the number of studies the estimates are based on. *Italic* entries are not significant in the ANOVA at a 5% level, but are included, like the overall mean, to show otherwise interesting or some selected groups. The <u>underlined</u> entry is the reference category.

on public sites. Since it benefits all, it is also common to store a study's reference in various data bases to make it available to a broader audience[1]. To a large extent this is not the case for many

[1]Though acquiring a new study is almost no problem anymore, it is still difficult to make a study known to the relevant audience when it is not published in one of the most popular journals.

old studies which aren't published in a well known journal; especially working papers and articles of minor journals. Although some data bases try to index such papers, their success is limited because many papers are not available anymore. Whereas an electronic format (in most cases pdf) stores the data in an easy accessible (text-)form, many old studies are, if at all, only available in a (scanned) paper-form or on micro-film. These limitations of the available data bases were proved to be strong in our research process, as can be seen by the ratio between the number of references retrieved by the data bases and the references based on citations only [2], as depicted in figures A.3 and A.4.

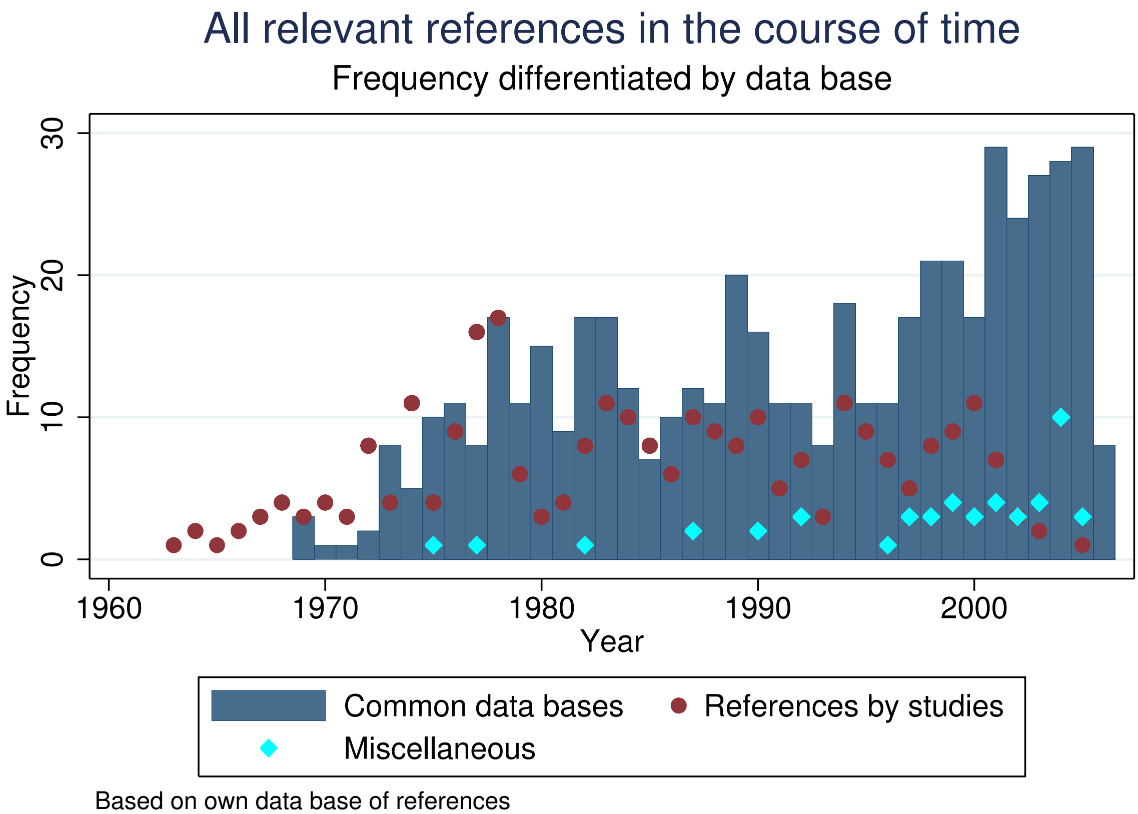
Figure A.3: Source of all 3598 references in the course of time (1950-2006)



After the first stage of the search process and attaining all available studies, 34.63% (31.89% of all journal articles) of all references are taken from citations of acquired studies and another 3.22% (5.75%) from miscellaneous sources (refer to figure A.3 for the distribution over time). This picture remains robust with respect to the relevance of these studies. In regard to all studies which were attained and judged relevant (see figure A.4), 32.86 (32.82%) are based on citations and 5.71% (2.25%) on other sources (numbers in parentheses are restricted to journal articles only). The drop in the beginning of the 21th century of the references-by-studies is partly based

[2]Retrieval by data bases also included those studies which cite Becker (1968) and Ehrlich (1973), as described in section 3.1.

Figure A.4: Source of all 840 relevant references in the course of time (1960-2006)



on the fact that studies require some time to be referenced by other studies and the newest acquired studies are from the first half of 2006.

A.2 Publication Bias

In the case of no effect the expected value of α_0 (see subsection 3.4.1) is computed as

$$
\begin{aligned}
\mathbb{E}[\log(|t|)] &= 2\int_0^\infty \log(t)\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{1}{\nu}\right)}\left(\frac{\nu+t^2}{\nu}\right)^{-\frac{\nu+1}{2}} dt \\
&= \frac{-\log\left(\frac{1}{\nu}\right)-\Psi\left(\frac{\nu}{2}\right)-\gamma-2\log 2}{2} \\
&= -\frac{\log 2+\gamma}{2}+\frac{\log\nu-\Psi(\nu/2)-\log 2}{2} \\
&= -\frac{\log 2+\gamma}{2}+O(1/\nu) \\
&\approx -0.6351814
\end{aligned}
$$

The DiGamma-function Ψ is $\frac{d\Gamma(x)}{dx}\frac{1}{\Gamma(x)}$ and the Euler-constant γ is approximately 0.5772157. The term $\frac{\log\nu-\Psi(0.5\nu)-\log 2}{2}$ is $O(1/\nu)$ because

$$\limsup_{\nu\to\infty}\left|\nu\frac{\log\nu-\Psi(0.5\nu)-\log 2}{2}\right| = 0.5.$$

Indeed, $-\frac{\log 2+\gamma}{2}$ is the result derived from the standard normal distribution.

B Statistics of Variables and Studies

There is no safety in numbers, or in anything else.

James Thurber

The Codebook contains all variables in our data base, including the variables we have discarded for various reasons (mainly due to the lack of observations). Afterwards, we show all studies in our meta-data base along with some basic information (studied crime types, overall opinion of the authors and the user who processed the study) and its reference (including the year of publication).

B.1 Codebook

The variables are sorted by their relationship to study- and estimate-properties and whether they are metric or non-metric. Only non-empty entries of a variable are listed.

Table B.1: Descriptive statistics of the non-metric study-variables

	Unweighted			Weighted		
Variable	N	Percent	Cum.	N	Percent	Cum.
Study: user						
ah	122	17.43	17.43	1209.27	18.52	18.52
aw	197	28.14	45.57	1911.78	29.28	47.80
kr	22	3.14	48.71	175.17	2.68	50.48
mw	9	1.29	50.00	88.76	1.36	51.84
tr	350	50.00	100.00	3145.01	48.16	100.00
Total	700	100.00		6530.00	100.00	
Study: publication, type						
conference paper	11	1.57	1.57	98.30	1.51	1.51
journal	604	86.29	87.86	5634.98	86.29	87.80
edited volume	28	4.00	91.86	260.70	3.99	91.79
working paper	34	4.86	96.71	321.90	4.93	96.72
miscellaneous	6	0.86	97.57	54.32	0.83	97.55
book	17	2.43	100.00	159.80	2.45	100.00
Total	700	100.00		6530.00	100.00	
Study: publication, thesis						
PhD thesis	16	2.29	2.29	149.61	2.29	2.29
none	684	97.71	100.00	6380.39	97.71	100.00
Total	700	100.00		6530.00	100.00	
Study: publication, country						
Argentina	3	0.43	0.43	30.55	0.47	0.47

. . . last page of table B.1 continued

Variable	Unweighted N	Unweighted Percent	Unweighted Cum.	Weighted N	Weighted Percent	Weighted Cum.
Australia	4	0.57	1.00	40.74	0.62	1.09
Belgium	2	0.29	1.29	20.37	0.31	1.40
Canada	23	3.29	4.57	234.24	3.59	4.99
Chile	1	0.14	4.71	10.18	0.16	5.15
China	1	0.14	4.86	0.00	0.00	5.15
Columbia	1	0.14	5.00	6.79	0.10	5.25
France	1	0.14	5.14	0.00	0.00	5.25
Germany	31	4.43	9.57	298.84	4.58	9.83
Ireland	2	0.29	9.86	20.37	0.31	10.14
Italy	1	0.14	10.00	10.18	0.16	10.30
Republic of Korea	1	0.14	10.14	10.18	0.16	10.45
Netherlands	18	2.57	12.71	166.34	2.55	13.00
Russia	1	0.14	12.86	5.09	0.08	13.08
Sweden	3	0.43	13.29	25.46	0.39	13.47
Switzerland	1	0.14	13.43	10.18	0.16	13.62
United Kingdom	56	8.00	21.43	539.47	8.26	21.88
USA	550	78.57	100.00	5101.00	78.12	100.00
Total	700	100.00		6530.00	100.00	
Study: publication, discipline						
missing	12	1.71	1.71	104.50	1.60	1.60
law	35	5.00	6.71	346.27	5.30	6.90
criminology	142	20.29	27.00	1335.35	20.45	27.35
economy	250	35.71	62.71	2348.48	35.96	63.32
sociology	135	19.29	82.00	1267.33	19.41	82.72
psychology	20	2.86	84.86	183.32	2.81	85.53
miscellaneous	103	14.71	99.57	919.29	14.08	99.61
not applicable	3	0.43	100.00	25.46	0.39	100.00
Total	700	100.00		6530.00	100.00	
Study: author, country						
missing	5	0.68	0.68	50.92	0.75	0.75
Argentina	3	0.41	1.09	30.55	0.45	1.20
Australia	16	2.19	3.28	132.40	1.94	3.14
Brasilia	4	0.55	3.83	37.34	0.55	3.69
Canada	34	4.65	8.48	336.08	4.93	8.62
Chile	1	0.14	8.62	6.79	0.10	8,72
China	4	0.55	9.17	40.74	0.60	9.32
Democratic Republic of the Congo	1	0.14	9.30	10.18	0.15	9.47
Finland	6	0.82	10.12	49.65	0.73	10.20
France	1	0.14	10.26	10.18	0.15	10.35
Germany	22	3.01	13.27	207.58	3.05	13.39
India	2	0.27	13.54	20.37	0.30	13.69
Ireland	2	0.27	13.82	20.37	0.30	13.99
Israel	9	1.23	15.05	84.87	1.25	15.24
Italy	3	0.41	15.46	30.55	0.45	15.69
Japan	1	0.14	15.60	0.00	0.00	15.69

. . . last page of table B.1 continued

Variable	Unweighted N	Unweighted Percent	Unweighted Cum.	Weighted N	Weighted Percent	Weighted Cum.
Republic of Korea	3	0.41	16.01	30.55	0.45	16.14
Netherlands	8	1.09	17.10	81.47	1.20	17.33
New Zealand	2	0.27	17.37	20.37	0.30	17.63
Norway	2	0.27	17.65	19.69	0.29	17.82
Russia	1	0.14	17.78	5.09	0.07	17.99
Singapore	2	0.27	18.06	10.18	0.15	18.14
Spain	2	0.27	18.33	20.37	0.30	18.44
Sweden	6	0.82	19.15	61.11	0.90	19.34
Switzerland	8	1.09	20.25	66.20	0.97	20.31
Turkey	1	0.14	20.38	10.18	0.15	20.46
United Kingdom	30	4.10	24.49	275.35	4.04	24.50
USA	552	75.51	100.00	5142.61	75.50	100.00
Total	731	100.00		6811.77	100.00	
Study: author, discipline						
missing	54	7.28	7.28	429.44	6.19	6.19
law	26	3.50	10.78	233.14	3.36	9.55
criminology	75	10.11	20.89	738.41	10.64	20.19
economy	296	39.89	60.78	2807.36	40.45	60.63
sociology	178	23.99	84.77	1689.98	24.35	84.98
psychology	29	3.91	88.68	274.98	3.96	88.95
miscellaneous	84	11.32	100.00	767.25	11.05	100.00
Total	742	100.00		6940.56	100.00	
Study: author, institution						
missing	21	3.00	3.00	183.32	2.81	2.81
law	28	4.00	7.00	253.21	3.88	6.68
criminology	81	11.57	18.57	784.08	12.01	18.69
economy	287	41.00	59.57	2728.86	41.79	60.48
sociology	149	21.29	80.86	1390.23	21.29	81.77
psychology	24	3.43	84.29	218.96	3.35	85.13
miscellaneous	110	15.71	100.00	971.33	14.87	100.00
Total	700	100.00		6530.00	100.00	
Study: type						
explorative	37	5.29	5.29	264.01	4.04	4.04
not explorative	663	94.71	100.00	6265.99	95.96	100.00
Total	700	100.00		6530.00	100.00	
Study: theory						
traditional	349	49.86	49.86	3374.81	51.68	51.68
economical	269	38.43	88.29	2465.96	37.76	89.45
miscellaneous	82	11.71	100.00	689.24	10.55	100.00
Total	700	100.00		6530.00	100.00	
Study: time series						
no	458	65.43	65.43	4294.65	65.77	65.77
yes	242	34.57	100.00	2235.35	34.23	100.00
Total	700	100.00		6530.00	100.00	

. . . last page of table B.1 continued

Variable	Unweighted N	Unweighted Percent	Unweighted Cum.	Weighted N	Weighted Percent	Weighted Cum.
Study: cross section						
no	534	76.29	76.29	5012.89	76.77	76.77
yes	166	23.71	100.00	1517.11	23.23	100.00
Total	700	100.00		6530.00	100.00	
Study: panel data						
no	598	85.43	85.43	5530.25	84.69	84.69
yes	102	14.57	100.00	999.75	15.31	100.00
Total	700	100.00		6530.00	100.00	
Study: single survey						
no	577	82.43	82.43	5368.70	82.22	82.22
yes	123	17.57	100.00	1161.30	17.78	100.00
Total	700	100.00		6530.00	100.00	
Study: repeated survey						
no	674	96.29	96.29	6325.37	96.87	96.87
yes	26	3.71	100.00	204.63	3.13	100.00
Total	700	100.00		6530.00	100.00	
Study: panel survey						
no	662	94.57	94.57	6162.95	94.38	94.38
yes	38	5.43	100.00	367.05	5.62	100.00
Total	700	100.00		6530.00	100.00	
Study: experiment (laboratory)						
no	671	95.86	95.86	6244.84	95.63	95.63
yes	29	4.14	100.00	285.16	4.37	100.00
Total	700	100.00		6530.00	100.00	
Study: experiment (field, researcher initiative)						
no	690	98.57	98.57	6428.16	98.44	98.44
yes	10	1.43	100.00	101.84	1.56	100.00
Total	700	100.00		6530.00	100.00	
Study: experiment (field, institutional initiative)						
no	689	98.43	98.43	6428.16	98.44	98.44
yes	11	1.57	100.00	101.84	1.56	100.00
Total	700	100.00		6530.00	100.00	
Study: experiment (natural)						
no	679	97.00	97.00	6333.10	96.98	96.98
yes	21	3.00	100.00	196.90	3.02	100.00
Total	700	100.00		6530.00	100.00	
Study: document analysis						
no	694	99.14	99.14	6468.89	99.06	99.06
yes	6	0.86	100.00	61.11	0.94	100.00
Total	700	100.00		6530.00	100.00	
Study: multiple dimensions						
no	700	100.00	100.00	6530.00	100.00	100.00
Total	700	100.00		6530.00	100.00	
Study: miscellaneous						
no	698	99.71	99.71	6509.63	99.69	99.69

. . . last page of table B.1 continued

	Unweighted			Weighted		
Variable	N	Percent	Cum.	N	Percent	Cum.
yes	2	0.29	100.00	20.37	0.31	100.00
Total	700	100.00		6530.00	100.00	
Study: experimental						
missing	2	0.29	0.29	20.37	0.31	0.31
experimental	47	6.71	7.00	445.57	6.82	7.14
quasi experimental	65	9.29	16.29	572.87	8.77	15.91
not experimental	586	83.71	100.00	491.20	84.09	100.00
Total	700	100.00		6530.00	100.00	
Study: first population						
missing	8	1.11	1.11	72.99	1.09	1.09
Argentina	4	0.56	1.67	40.74	0.61	1.69
Australia	13	1.81	3.48	112.03	1.67	3.36
Austria	1	0.14	3.62	10.18	0.15	3.51
Bangladesh	1	0.14	3.76	10.18	0.15	3.67
Brazil	1	0.14	3.90	10.18	0.15	3.82
Canada	28	3.90	7.80	280.07	4.17	7.99
China	3	0.42	8.22	30.55	0.46	8.44
Ecuador	1	0.14	8.36	10.18	0.15	8.60
Egypt	1	0.14	8.50	10.18	0.15	8.75
Finland	7	0.97	9.47	59.83	0.89	9.64
France	3	0.42	9.89	26.19	0.39	10.03
Germany	19	2.65	12.53	177.03	2.64	12.67
India	1	0.14	12.67	0.00	0.00	12.67
Ireland	2	0.28	12.95	20.37	0.30	12.97
Israel	4	0.56	13.51	33.95	0.51	13.48
Italy	4	0.56	14.07	40.74	0.61	14.08
Jamaica	1	0.14	14.21	10.18	0.15	14.23
Japan	4	0.56	14.76	30.55	0.46	14.69
Republic of Korea	2	0.28	15.04	20.37	0.30	14.99
Malaysia	2	0.28	15.32	20.37	0.30	15.30
Netherlands	8	1.11	16.43	81.47	1.21	16.51
New Zealand	4	0.56	16.99	40.74	0.61	17.12
Norway	3	0.42	17.41	24.78	0.37	17.49
Russia	2	0.28	17.69	15.28	0.23	17.71
South Africa	1	0.14	17.83	10.18	0.15	17.86
Spain	1	0.14	17.97	10.18	0.15	18.02
Sweden	9	1.25	19.22	91.66	1.37	19.38
Switzerland	8	1.11	20.33	66.20	0.99	20.37
Turkey	1	0.14	20.47	10.18	0.15	20.52
United Kingdom	33	4.60	25.07	327.26	4.87	25.39
USA	538	74.83	100.00	5008.50	74.61	100.00
Total	718	100.00				
Study: second population						
Albania	1	3.03	3.03	14.94	3.48	3.48
Finland	1	3.03	6.06	14.94	3.48	6.97

. . . last page of table B.1 continued

Variable	Unweighted N	Percent	Cum.	Weighted N	Percent	Cum.
Germany	2	6.06	12.12	29.89	6.97	13.93
Israel	1	3.03	15.15	14.94	3.48	17.42
Norway	2	6.06	21.21	29.89	6.97	24.38
Sweden	1	3.03	24.24	7.47	1.74	26.13
USA	25	75.76	100.00	16.92	73.87	100.00
Total	33	100.00		429.00	100.00	
Study: sample base, first population						
missing	3	0.43	0.43	30.55	0.47	0.47
complete country	262	37.43	37.86	2410.25	36.91	37.38
partial country	435	62.14	100.00	4089.20	62.62	100.00
Total	700	100.00		6530.00	100.00	
Study: sample base, second population						
missing	667	95.29	95.29	6233.82	95.46	95.46
complete country	16	2.29	97.57	140.94	2.16	97.62
partial country	17	2.43	100.00	155.24	2.38	100.00
Total	700	100.00		6530.00	100.00	
Study: sample unit, first population, nation						
no	587	83.86	83.86	5509.41	84.37	84.37
yes	113	16.14	100.00	1020.59	15.63	100.00
Total	700	100.00		6530.00	100.00	
Study: sample unit, first population, states						
no	545	77.86	77.86	5099.55	78.09	78.09
yes	155	22.14	100.00	1430.45	21.91	100.00
Total	700	100.00		6530.00	100.00	
Study: sample unit, first population, counties						
no	525	75.00	75.00	4887.24	74.84	74.84
yes	175	25.00	100.00	1642.76	25.16	100.00
Total	700	100.00		6530.00	100.00	
Study: sample unit, first population, individuals						
no	570	81.43	81.43	5332.31	81.66	81.66
yes	130	18.57	100.00	1197.69	18.34	100.00
Total	700	100.00		6530.00	100.00	
Study: sample unit, first population, groups						
no	612	87.43	87.43	5663.11	86.72	86.72
yes	88	12.57	100.00	866.89	13.28	100.00
Total	700	100.00		6530.00	100.00	
Study: sample unit, first population, actions						
no	696	99.43	99.43	6489.26	99.38	99.38
yes	4	0.57	100.00	40.74	0.62	100.00
Total	700	100.00		6530.00	100.00	
Study: sample unit, first population, miscellaneous						
no	650	92.86	92.86	6046.35	92.59	92.59
yes	50	7.14	100.00	483.65	7.41	100.00
Total	700	100.00		6530.00	100.00	
Study: sample unit, second population, nation						

...last page of table B.1 continued

Variable	Unweighted N	Unweighted Percent	Unweighted Cum.	Weighted N	Weighted Percent	Weighted Cum.
no	697	99.57	99.57	6500.72	99.55	99.55
yes	3	0.43	100.00	29.28	0.45	100.00
Total	700	100.00		6530.00	100.00	
Study: sample unit, second population, states						
no	695	99.29	99.29	6481.99	99.26	99.26
yes	5	0.71	100.00	48.01	0.74	100.00
Total	700	100.00		6530.00	100.00	
Study: sample unit, second population, counties						
no	690	98.57	98.57	6450.90	98.79	98.79
yes	10	1.43	100.00	79.10	1.21	100.00
Total	700	100.00		6530.00	100.00	
Study: sample unit, second population, individuals						
no	688	98.29	98.29	6420.52	98.32	98.32
yes	12	1.71	100.00	109.48	1.68	100.00
Total	700	100.00		6530.00	100.00	
Study: sample unit, second population, groups						
no	697	99.57	99.57	6499.69	99.54	99.54
yes	3	0.43	100.00	30.31	0.46	100.00
Total	700	100.00		6530.00	100.00	
Study: sample unit, second population, actions						
no	700	100.00	100.00	6530.00	100.00	100.00
Total	700	100.00		6530.00	100.00	
Study: sample unit, second population, miscellaneous						
no	699	99.86	99.86	6519.82	99.84	99.84
yes	1	0.14	100.00	10.18	0.16	100.00
Total	700	100.00		6530.00	100.00	
Study: sample individuals, first population						
missing	4	0.57	0.57	40.74	0.62	0.62
population	455	65.00	65.57	4165.76	63.79	64.42
students	78	11.14	76.71	762.06	11.67	76.09
pupils	23	3.29	80.00	200.49	3.07	79.16
prisoners	4	0.57	80.57	40.74	0.62	79.78
miscellaneous	136	19.43	100.00	1320.21	20.22	100.00
Total	700	100.00		6530.00	100.00	
Study: sample individuals, second population						
missing	668	95.43	95.43	6244.00	95.62	95.62
population	19	2.71	98.14	166.57	2.55	98.17
students	3	0.43	98.57	30.55	0.47	98.64
pupils	1	0.14	98.71	9.94	0.15	98.79
miscellaneous	9	1.29	100.00	78.93	1.21	100.00
Total	700	100.00		6530.00	100.00	
Study: type of first sample						
missing	413	59.00	59.00	3850.93	58.97	58.97
complete	74	10.57	69.57	641.15	9.82	68.79
random	118	16.86	86.43	1131.75	17.33	86.12

. . . last page of table B.1 continued

Variable	Unweighted N	Percent	Cum.	Weighted N	Percent	Cum.
quota	16	2.29	88.71	152.77	2.34	88.46
extreme	3	0.43	89.14	30.55	0.47	88.93
unsystematic	76	10.86	100.00	722.85	11.07	100.00
Total	700	100.00		6530.00	100.00	
Study: type of second sample						
missing	685	97.86	97.86	6397.68	97.97	97.97
complete	2	0.29	98.14	17.65	0.27	98.24
random	10	1.43	99.57	90.91	1.39	99.64
extreme	1	0.14	99.71	10.18	0.16	99.79
unsystematic	2	0.29	100.00	13.58	0.21	100.00
Total	700	100.00		6530.00	100.00	
Study: data base						
missing	63	9.00	9.00	629.50	9.64	9.64
PKS	8	1.14	10.14	74.93	1.15	10.79
UCR	163	23.29	33.43	1494.18	22.88	33.67
miscellaneous	280	40.00	73.43	2617.58	40.09	73.75
not public	186	26.57	100.00	1713.81	26.25	100.00
Total	700	100.00		6530.00	100.00	
Study: income, first population						
missing	628	89.71	89.71	5815.92	89.06	89.06
above average	4	0.57	90.29	40.74	0.62	89.69
representative	7	1.00	91.29	68.74	1.05	90.74
below average	61	8.71	100.00	604.60	9.26	100.00
Total	700	100.00		6530.00	100.00	
Study: income, second population						
missing	695	99.29	99.29	6481.87	99.26	99.26
representative	1	0.14	99.43	7.64	0.12	99.38
below average	4	0.57	100.00	40.49	0.62	100.00
Total	700	100.00		6530.00	100.00	
Study: education, first population						
missing	630	90.00	90.00	5837.39	89.39	89.39
above average	63	9.00	99.00	624.96	9.57	98.96
representative	6	0.86	99.86	58.56	0.90	99.86
below average	1	0.14	100.00	9.09	0.14	100.00
Total	700	100.00		6530.00	100.00	
Study: education, second population						
missing	697	99.57	99.57	6501.99	99.57	99.57
above average	2	0.29	99.86	20.37	0.31	99.88
representative	1	0.14	100.00	7.64	0.12	100.00
Total	700	100.00		6530.00	100.00	
Study: class, first population						
missing	683	97.57	97.57	6361.91	97.43	97.43
upper above average	4	0.57	98.14	40.74	0.62	98.05
no class above average	7	1.00	99.14	68.74	1.05	99.10
lower above average	6	0.86	100.00	58.61	0.90	100.00

. . . last page of table B.1 continued

	Unweighted			Weighted		
Variable	N	Percent	Cum.	N	Percent	Cum.
Total	700	100.00		6530.00	100.00	
Study: class, second population						
missing	698	99.71	99.71	6512.18	99.73	99.73
no class above average	1	0.14	99.86	7.64	0.12	99.84
lower above average	1	0.14	100.00	10.18	0.16	100.00
Total	700	100.00		6530.00	100.00	
Study: disadvantaged group, first population						
missing	686	98.00	98.00	6399.00	97.99	97.99
yes	10	1.43	99.43	90.26	1.38	99.38
no	4	0.57	100.00	40.74	0.62	100.00
Total	700	100.00		6530.00	100.00	
Study: disadvantaged group, second population						
missing	700	100.00	100.00	6530.00	100.00	100.00
Total	700	100.00		6530.00	100.00	
Study: percentage convicted, first population						
missing	692	98.86	98.86	6449.62	98.77	98.77
1-25 percent	1	0.14	99.00	10.18	0.16	98.93
26-50 percent	1	0.14	99.14	9.09	0.14	99.06
51-75 percent	2	0.29	99.43	20.37	0.31	99.38
76-100 percent	4	0.57	100.00	40.74	0.62	100.00
Total	700	100.00		6530.00	100.00	
Study: percentage convicted, second population						
missing	700	100.00	100.00	6530.00	100.00	100.00
Total	700	100.00		6530.00	100.00	
Study: main location, first population						
missing	655	93.57	93.57	6100.98	93.43	93.43
1000-5000 inhabitants	1	0.14	93.71	10.18	0.16	93.59
5000-20000 inhabitants	2	0.29	94.00	19.27	0.30	93.88
20000-100000 inhabitants	6	0.86	94.86	50.92	0.78	94.66
100000-500000 inhabitants	9	1.29	96.14	91.66	1.40	96.06
>500000 inhabitants	24	3.43	99.57	226.43	3.47	99.53
mixed	3	0.43	100.00	30.55	0.47	100.00
Total	700	100.00		6530.00	100.00	
Study: main location, second population						
missing	698	99.71	99.71	6512.35	99.73	99.73
>500000 inhabitants	2	0.29	100.00	17.65	0.27	100.00
Total	700	100.00		6530.00	100.00	
Study: urbanity, first population						
missing	644	92.00	92.00	5986.11	91.67	91.67
smaller cities above average	7	1.00	93.00	68.92	1.06	92.73
representative	7	1.00	94.00	68.74	1.05	93.78
large cities above average	42	6.00	100.00	406.23	6.22	100.00
Total	700	100.00		6530.00	100.00	
Study: urbanity, second population						
missing	698	99.71	99.71	6512.35	99.73	99.73

...last page of table B.1 continued

Variable	Unweighted N	Unweighted Percent	Unweighted Cum.	Weighted N	Weighted Percent	Weighted Cum.
large cities above average	2	0.29	100.00	17.65	0.27	100.00
Total	700	100.00		6530.00	100.00	
Study: claims representativeness						
missing	334	47.71	47.71	3139.46	48.08	48.08
yes	137	19.57	67.29	1266.48	19.39	67.47
no	229	32.71	100.00	2124.06	32.53	100.00
Total	700	100.00		6530.00	100.00	
Study: tests representativeness						
missing	161	23.00	23.00	1566.28	23.99	23.99
yes	21	3.00	26.00	186.71	2.86	26.85
no	518	74.00	100.00	4777.01	73.15	100.00
Total	700	100.00		6530.00	100.00	
Study: data are representative						
missing	681	97.29	97.29	6357.55	97.36	97.36
yes	3	0.43	97.71	28.76	0.44	97.80
no	16	2.29	100.00	143.70	2.20	100.00
Total	700	100.00		6530.00	100.00	
Study: performs pretest						
yes	19	2.71	2.71	168.48	2.58	2.58
no	681	97.29	100.00	6361.52	97.42	100.00
Total	700	100.00		6530.00	100.00	
Study: pretest questions						
missing	38	5.43	5.43	340.39	5.21	5.21
open	10	1.43	6.86	91.66	1.40	6.62
open and closed	15	2.14	9.00	134.09	2.05	8.67
closed	148	21.14	30.14	1404.47	21.51	30.18
not applicable	489	69.86	100.00	4559.39	69.82	100.00
Total	700	100.00		6530.00	100.00	
Study: error and plausibility checks						
yes	220	31.43	31.43	2053.09	31.44	31.44
no	480	68.57	100.00	4476.91	68.56	100.00
Total	700	100.00		6530.00	100.00	
Study: check for reliability						
yes	33	4.71	4.71	324.26	4.97	4.97
no	667	95.29	100.00	6205.74	95.03	100.00
Total	700	100.00		6530.00	100.00	
Study: reliability check results						
missing	669	95.57	95.57	6215.93	95.19	95.19
for all variables	9	1.29	96.86	90.26	1.38	96.57
for some variables	22	3.14	100.00	223.81	3.43	100.00
Total	700	100.00		6530.00	100.00	
Study: values of reliability checks						
missing	669	95.57	95.57	6226.11	95.35	95.35
yes	25	3.57	99.14	252.97	3.87	99.22
no	6	0.86	100.00	50.92	0.78	100.00

. . . last page of table B.1 continued

Variable	Unweighted N	Unweighted Percent	Unweighted Cum.	Weighted N	Weighted Percent	Weighted Cum.
Total	700	100.00		6530.00	100.00	
Study: reliability method						
missing	673	96.14	96.14	6256.67	95.81	95.81
Cronbach's α	23	3.29	99.43	232.60	3.56	99.38
Guttman	1	0.14	99.57	10.18	0.16	99.53
Correlation	1	0.14	99.71	10.18	0.16	99.69
Miscellaneous	2	0.29	100.00	20.37	0.31	100.00
Total	700	100.00		6530.00	100.00	
Study: reliable variables						
missing	668	95.43	95.43	6205.74	95.03	95.03
some	2	0.29	95.71	20.37	0.31	95.35
yes	25	3.57	99.29	254.37	3.90	99.24
unknown	5	0.71	100.00	49.52	0.76	100.00
Total	700	100.00		6530.00	100.00	
Study: check for validity						
yes	17	2.43	2.43	136.53	2.09	2.09
no	683	97.57	100.00	6393.47	97.91	100.00
Total	700	100.00		6530.00	100.00	
Study: validity test						
missing	686	98.00	98.00	6413.84	98.22	98.22
for all variables	2	0.29	98.29	15.28	0.23	98.46
for some variables	12	1.71	100.00	100.89	1.54	100.00
Total	700	100.00		6530.00	100.00	
Study: validity test method						
missing	687	98.14	98.14	6424.02	98.38	98.38
by criteria	7	1.00	99.14	62.46	0.96	99.33
by construction	3	0.43	99.57	18.05	0.28	99.61
miscellaneous	3	0.43	100.00	25.46	0.39	100.00
Total	700	100.00		6530.00	100.00	
Study: valid variables						
missing	683	97.57	97.57	6393.47	97.91	97.91
no	1	0.14	97.71	10.18	0.16	98.07
yes	13	1.86	99.57	105.98	1.62	99.69
unknown	3	0.43	100.00	20.37	0.31	100.00
Total	700	100.00		6530.00	100.00	
Study: conditions for significance check fulfilled						
yes	348	49.71	49.71	3317.27	50.80	50.80
almost	274	39.14	88.86	2527.22	38.70	89.50
almost not	52	7.43	96.29	451.27	6.91	96.41
no	26	3.71	100.00	234.24	3.59	100.00
Total	700	100.00		6530.00	100.00	
Study: tests of significance						
missing	3	0.43	0.43	30.55	0.47	0.47
yes	614	87.71	88.14	5915.06	90.58	91.05
no	83	11.86	100.00	584.39	8.95	100.00

...last page of table B.1 continued

Variable	Unweighted N	Unweighted Percent	Unweighted Cum.	Weighted N	Weighted Percent	Weighted Cum.
Total	700	100.00		6530.00	100.00	
Study: values of relationships						
missing	3	0.43	0.43	30.55	0.47	0.47
yes	619	88.43	88.86	5913.85	90.56	91.03
no	78	11.14	100.00	585.60	8.97	100.00
Total	700	100.00		6530.00	100.00	
Study: uses covariates						
yes	553	79.00	79.00	5260.02	80.55	80.55
no	147	21.00	100.00	1269.98	19.45	100.00
Total	700	100.00		6530.00	100.00	
Study: problems reported by author						
yes	238	34.00	34.00	2194.16	33.60	33.60
no	462	66.00	100.00	4335.84	66.40	100.00
Total	700	100.00		6530.00	100.00	
Study: problems reported by reader						
no	189	27.00	27.00	1839.95	28.18	28.18
some	438	62.57	89.57	4151.97	63.58	91.76
severe	73	10.43	100.00	538.07	8.24	100.00
Total	700	100.00		6530.00	100.00	
Study: quality index						
0 (good)	138	19.71	19.71	1347.68	20.64	20.64
1	2	0.29	20.00	20.37	0.31	20.95
2	49	7.00	27.00	471.90	7.23	28.18
3	281	40.14	67.14	2687.72	41.16	69.34
4	28	4.00	71.14	254.70	3.90	73.24
5	129	18,43	89.57	1209.55	18.52	91.76
6	43	6.14	95.71	300.44	4.60	96.36
7	2	0.29	96.00	10.18	0.16	96.52
8 (bad)	28	4.00	100.00	227.45	3.48	100.00
Total	700	100.00		6530.00	100.00	
Study: author opinion, violent crime, probability						
strongly agree	69	20.97	20.97	871.04	21.47	21.47
partially agree	101	30.70	51.67	1259.70	31.05	52.52
indecisive	41	12.46	64.13	462.15	11.39	63.91
partially disagree	67	20.36	84.50	834.94	20.58	84.49
fully disagree	51	15.50	100.00	629.17	15.51	100.00
Total	329	100.00		4057.00	100.00	
Study: author opinion, violent crime, severity						
strongly agree	32	20.38	20.38	489.37	21.93	21.93
partially agree	37	23.57	43.95	550.24	24.65	46.58
indecisive	29	18.47	62.42	366.75	16.43	63.01
partially disagree	30	19.11	81.53	398.02	17.83	80.84
fully disagree	29	18.47	100.00	427.61	19.16	100.00
Total	157	100.00		2232.00	100.00	
Study: author opinion, property crime, probability						

...last page of table B.1 continued

Variable	Unweighted N	Unweighted Percent	Unweighted Cum.	Weighted N	Weighted Percent	Weighted Cum.
strongly agree	106	31.74	31.74	1391.31	32.19	32.19
partially agree	109	32.63	64.37	1439.27	33.30	65.49
indecisive	37	11.08	75.45	462.43	10.70	76.19
partially disagree	48	14.37	89.82	585.60	13.55	89.75
fully disagree	34	10.18	100.00	443.19	10.25	100.00
Total	334	100.00		4322.00	100.00	
Study: author opinion, property crime, severity						
strongly agree	29	19.33	19.33	492.05	18.96	18.96
partially agree	31	20.67	40.00	571.18	22.01	40.97
indecisive	25	16.67	56.67	396.96	15.30	56.27
partially disagree	40	26.67	83.33	678.67	26.15	82.42
fully disagree	25	16.67	100.00	456.14	17.58	100.00
Total	150	100.00		2595.00	100.00	
Study: author opinion, other crime, probability						
strongly agree	43	16.04	16.04	375.61	15.09	15.09
partially agree	117	43.66	59.70	1136.10	45.65	60.74
indecisive	16	5.97	65.67	141.26	5.68	66.41
partially disagree	46	17.16	82.84	433.44	17.41	83.83
fully disagree	46	17.16	100.00	402.59	16.17	100.00
Total	268	100.00		2489.00	100.00	
Study: author opinion, other crime, severity						
strongly agree	23	14.29	14.29	266.95	14.60	14.60
partially agree	40	24.84	39.13	454.23	24.83	39.43
indecisive	18	11.18	50.31	211.10	11.54	50.97
partially disagree	32	19.88	70.19	369.92	20.23	71.20
fully disagree	48	29.81	100.00	526.80	28.80	100.00
Total	161	100.00		1829.00	100.00	

Columns 2-4 are not weighted and refer to one estimate per study. Columns 5-7 are weighted and relate to all valid estimates. For reasons of parsimony the authors and journals are not displayed in this list.

Table B.2: Descriptive statistics of the non-metric estimate-variables

Variable	Unweighted N	Unweighted Percent	Unweighted Cum.	Weighted N	Weighted Percent	Weighted Cum.
Estimate: deterrence is focus-variable						
yes	7133	91.19	91.19	5563.96	85.21	85.21
no	689	8.81	100.00	966.04	14.79	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: complete sample						
missing	1	0.01	0.01			
yes	6209	79.38	79.39	5558.02	85.12	85.12
no	1612	20.61	100.00	971.98	14.88	100.00
Total	7822	100.00		6530.00	100.00	

. . . last page of table B.2 continued

Variable	Unweighted N	Unweighted Percent	Unweighted Cum.	Weighted N	Weighted Percent	Weighted Cum.
Estimate: sub-sample, sex						
missing	7637	97.63	97.63	6384.54	97.77	97.77
female	73	0.93	98.57	51.12	0.78	98.56
male	112	1.43	100.00	94.34	1.44	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: sub-sample, age						
missing	7525	96.20	96.20	6416.82	98.27	98.27
younger	181	2.31	98.52	83.38	1.28	99.54
older	116	1.48	100.00	29.79	0.46	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: sub-sample, urbanity						
missing	7759	99.19	99.19	6461.51	98.95	98.95
urban	54	0.69	99.88	53.86	0.82	99.78
rural	9	0.12	100.00	14.63	0.22	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: sub-sample, education						
missing	7819	99.96	99.96	6509.63	99.69	99.69
low	2	0.03	99.99	15.28	0.23	99.92
high	1	0.01	100.00	5.09	0.08	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: sub-sample, norm acceptance						
missing	7816	99.92	99.92	6525.93	99.94	99.94
low	3	0.04	99.96	2.04	0.03	99.97
high	3	0.04	100.00	2.04	0.03	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: sub-sample, expected probability of detection						
missing	7709	98.56	98.56	6473.72	99.14	99.14
low	33	0.42	98.98	27.68	0.42	99.56
medium	21	0.27	99.25	6.46	0.10	99.66
high	59	0.75	100.00	22.13	0.34	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, index						
missing	7484	95.68	95.68	6296.90	96.43	96.43
additive, unweighted	69	0.88	96.56	41.47	0.64	97.07
additive, weighted	8	0.10	96.66	14.00	0.21	97.28
mean	1	0.01	96.68	0.68	0.01	97.29
multiplicative	173	2.21	98.89	125.79	1.93	99.22
miscellaneous	87	1.11	100.00	51.15	0.78	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, index items processed						
missing	7503	95.92	95.92	6323.47	96.84	96.84
raw	304	3.89	99.81	181.07	2.77	99.61
standardized	10	0.13	99.94	10.18	0.16	99.77
miscellaneous	5	0.06	100.00	15.28	0.23	100.00
Total	7822	100.00		6530.00	100.00	

. . . last page of table B.2 continued

Variable	Unweighted N	Unweighted Percent	Unweighted Cum.	Weighted N	Weighted Percent	Weighted Cum.
Estimate: study type						
Death Penalty	842	10.76	10.76	534.20	8.18	8.18
Crime Data	4066	51.98	62.75	3569.17	54.66	62.84
Survey	2569	32.84	95.59	1595.07	24.43	87.27
Experiment	345	4.41	100.00	831.57	12.73	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, death penalty, existence of death penalty						
no	7747	99.04	99.04	6463.02	98.97	98.97
yes	75	0.96	100.00	66.98	1.03	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, death penalty, conviction rate						
no	7815	99.91	99.91	6510.99	99.71	99.71
yes	7	0.09	100.00	19.01	0.29	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, death penalty, percentage of all convictions						
no	7726	98.77	98.77	6484.80	99.31	99.31
yes	96	1.23	100.00	45.20	0.69	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, death penalty, execution rate						
no	7288	93.17	93.17	6215.23	95.18	95.18
yes	534	6.83	100.00	314.77	4.82	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, death penalty, other						
no	7686	98.26	98.26	6433.28	98.52	98.52
yes	136	1.74	100.00	96.72	1.48	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, crime data, clearance rate						
no	7432	95.01	95.01	6186.25	94.74	94.74
yes	390	4.99	100.00	343.75	5.26	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, crime data, arrest rate						
no	6980	89.24	89.24	5911.20	90.52	90.52
yes	842	10.76	100.00	618.80	9.48	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, crime data, conviction rate						
no	7512	96.04	96.04	6275.12	96.10	96.10
yes	310	3.96	100.00	254.88	3.90	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, crime data, percentage of conviction, adult criminal law						
no	7822	100.00	100.00	6530.00	100.00	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, crime data, discontinuation rate						
no	7822	100.00	100.00	6530.00	100.00	100.00
Total	7822	100.00		6530.00	100.00	

. . . last page of table B.2 continued

Variable	Unweighted N	Unweighted Percent	Unweighted Cum.	Weighted N	Weighted Percent	Weighted Cum.
Estimate: exogenous, crime data, indictment rate						
no	7809	99.83	99.83	6523.82	99.91	99.91
yes	13	0.17	100.00	6.18	0.09	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, crime data, refraining from punishment (percentage of all convictions)						
no	7822	100.00	100.00	6530.00	100.00	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, crime data, parole rate						
no	7799	99.71	99.71	6519.58	99.84	99.84
yes	23	0.29	100.00	10.42	0.16	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, crime data, prison sentences per conviction						
no	7807	99.81	99.81	6517.90	99.81	99.81
yes	15	0.19	100.00	12.10	0.19	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, crime data, parole-rate (only if applicable)						
no	7822	100.00	100.00	6530.00	100.00	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, crime data, incarcerations (absolute or per capita)						
no	7720	98.70	98.70	6472.88	99.13	99.13
yes	102	1.30	100.00	57.12	0.87	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, crime data, incarceration rate						
no	7739	98.94	98.94	6488.14	99.36	99.36
yes	83	1.06	100.00	41.86	0.64	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, crime data, mean sentence length (sentenced)						
no	7549	96.51	96.51	6396.57	97.96	97.96
yes	273	3.49	100.00	133.43	2.04	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, crime data, inspections						
no	7744	99.00	99.00	6450.78	98.79	98.79
yes	78	1.00	100.00	79.22	1.21	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, crime data, mean sentence length (served)						
no	7611	97.30	97.30	6417.77	98.28	98.28
yes	211	2.70	100.00	112.23	1.72	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, crime data, time between offense and clearance						
no	7818	99.95	99.95	6527.45	99.96	99.96
yes	4	0.05	100.00	2.55	0.04	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, crime data, time between offense and conviction						
no	7820	99.97	99.97	6530.00	100.00	100.00

. . . last page of table B.2 continued

Variable	Unweighted N	Unweighted Percent	Unweighted Cum.	Weighted N	Weighted Percent	Weighted Cum.
yes	2	0.03	100.00	0.00	0.00	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, crime data, police expenditures						
no	7577	96.87	96.87	6272.29	96.05	96.05
yes	245	3.13	100.00	257.71	3.95	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, crime data, police strength						
no	7364	94.14	94.14	6016.77	92.14	92.14
yes	458	5.86	100.00	513.23	7.86	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, crime data, other						
no	7240	92.56	92.56	5972.23	91.46	91.46
yes	582	7.44	100.00	557.77	8.54	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, crime data, probability dummy (regime shift)						
no	7723	98.73	98.73	6301.50	96.50	96.50
yes	99	1.27	100.00	228.50	3.50	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, crime data, severity dummy (regime shift)						
no	7654	97.85	97.85	6295.68	96.41	96.41
yes	168	2.15	100.00	234.32	3.59	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, crime data, probation rate						
no	7787	99.55	99.55	6517.33	99.81	99.81
yes	35	0.45	100.00	12.67	0.19	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, crime data, incarceration per crime						
no	7728	98.80	98.80	6471.10	99.10	99.10
yes	94	1.20	100.00	58.90	0.90	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, crime data, convicted per crime						
no	7784	99.51	99.51	6489.16	99.37	99.37
yes	38	0.49	100.00	40.84	0.63	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, crime data, fine						
no	7774	99.40	99.40	6487.42	99.35	99.35
yes	47	0.60	100.00	42.58	0.65	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, survey, surveyed is . . .						
missing	5403	69.07	69.07	5003.23	76.62	76.62
delinquent	2159	27.60	96.68	1335.16	20.45	97.07
fictitious delinquent	260	3.32	100.00	191.61	2.93	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, survey, is experiment						
missing	5324	68.06	68.06	4961.81	75.98	75.98

...last page of table B.2 continued

Variable	Unweighted N	Unweighted Percent	Unweighted Cum.	Weighted N	Weighted Percent	Weighted Cum.
yes	27	0.35	68.41	59.83	0.92	76.90
no	2471	31.59	100.00	1508.36	23.10	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, survey, probability of detection by police						
no	7047	90.09	90.09	6063.72	92.86	92.86
yes	775	9.91	100.00	466.28	7.14	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, survey, probability of punishment by justice						
no	7338	93.81	93.81	6238.28	95.53	95.53
yes	484	6.19	100.00	291.72	4.47	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, survey, type of punishment						
no	7761	99.22	99.22	6512.80	99.74	99.74
yes	61	0.78	100.00	17.20	0.26	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, survey, severity of punishment by justice						
no	7307	93.42	93.42	6318.64	96.76	96.76
yes	515	6.58	100.00	211.36	3.24	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, survey, probability of civil punishment						
no	7818	99.95	99.95	6525.76	99.94	99.94
yes	4	0.05	100.00	4.24	0.06	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, survey, severity of civil punishment						
no	7821	99.99	99.99	6529.15	99.99	99.99
yes	1	0.01	100.00	0.85	0.01	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, survey, probability of punishment by employment law						
no	7798	99.69	99.69	6490.52	99.40	99.40
yes	24	0.31	100.00	39.48	0.60	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, survey, severity of punishment by employment law						
no	7808	99.82	99.82	6501.14	99.56	99.56
yes	14	0.18	100.00	28.86	0.44	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, survey, probability of other kind of punishment						
no	7775	99.40	99.40	6503.98	99.60	99.60
yes	47	0.60	100.00	26.02	0.40	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, survey, severity of other kind of punishment						
no	7809	99.83	99.83	6511.16	99.71	99.71
yes	13	0.17	100.00	18.84	0.29	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, survey, probability of detection by friends or family						
no	7789	99.58	99.58	6518.95	99.83	99.83

. . . last page of table B.2 continued

Variable	Unweighted N	Percent	Cum.	Weighted N	Percent	Cum.
yes	33	0.42	100.00	11.05	0.17	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, survey, probability of punishment by friends or family						
no	7583	96.94	96.94	6424.74	98.39	98.39
yes	239	3.06	100.00	105.26	1.61	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, survey, severity of punishment by friends or family						
no	7683	98.22	98.22	6467.95	99.05	99.05
yes	139	1.78	100.00	62.05	0.95	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, survey, probability of detection by others						
no	7779	99.45	99.45	6503.11	99.59	99.59
yes	43	0.55	100.00	26.89	0.41	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, survey, probability of punishment by others						
no	7751	99.09	99.09	6471.36	99.10	99.10
yes	71	0.91	100.00	58.64	0.90	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, survey, severity of punishment by others						
no	7803	99.76	99.76	6501.14	99.56	99.56
yes	19	0.24	100.00	28.86	0.44	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, survey, time between offense and clearance						
no	7795	99.65	99.65	6520.76	99.86	99.86
yes	27	0.35	100.00	9.24	0.14	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, survey, time between offense and conviction						
no	7812	99.87	99.87	6525.02	99.92	99.92
yes	10	0.13	100.00	4.98	0.08	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, survey, previous experience with police or justice						
no	7791	99.60	99.60	6473.99	99.14	99.14
yes	31	0.40	100.00	56.01	0.86	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, survey, other						
no	7653	97.84	97.84	6286.36	96.27	96.27
yes	169	2.16	100.00	243.64	3.73	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, survey, relates to . . .						
missing	5322	68.04	68.04	4960.67	75.97	75.97
the present	2378	30.40	98.44	1391.96	21.32	97.28
the past	122	1.56	100.00	177.38	2.72	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, experiment, is experiment						
missing	7477	95.59	95.59	5705.22	87.37	87.37

. . . last page of table B.2 continued

Variable	Unweighted N	Unweighted Percent	Unweighted Cum.	Weighted N	Weighted Percent	Weighted Cum.
yes	139	1.78	97.37	469.33	7.19	94.56
no	206	2.63	100.00	355.45	5.44	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, experiment, experimental variation of probability of detection						
no	7783	99.50	99.50	6394.04	97.92	97.92
yes	39	0.50	100.00	135.96	2.08	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, experiment, actual variation of probability of detection						
no	7697	98.40	98.40	6272.00	96.05	96.05
yes	125	1.60	100.00	258.00	3.95	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, experiment, actual variation of severity of punishment						
no	7779	99.45	99.45	6393.02	97.90	97.90
yes	43	0.55	100.00	136.98	2.10	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, experiment, experimental variation of severity of punishment						
no	7731	98.84	98.84	6390.29	97.86	97.86
yes	91	1.16	100.00	139.71	2.14	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, experiment, game losses (delinquent)						
no	7821	99.99	99.99	6524.91	99.92	99.92
yes	1	0.01	100.00	5.09	0.08	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, experiment, game losses (fictitious delinquent)						
no	7822	100.00	100.00	6530.00	100.00	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, experiment, utility balance (delinquent)						
no	7817	99.94	99.94	6509.63	99.69	99.69
yes	5	0.06	100.00	20.37	0.31	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, experiment, utility balance (fictitious delinquent)						
no	7822	100.00	100.00	6530.00	100.00	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, experiment, other						
no	7776	99.41	99.41	6389.12	97.84	97.84
yes	46	0.59	100.00	140.88	2.16	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, experiment, relates to . . .						
missing	7478	95.60	95.60	5593.19	85.65	85.65
present	328	4.19	99.80	906.25	13.88	99.53
past	16	0.20	100.00	30.55	0.47	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, relates to . . .						
missing	1413	18.06	18.06	825.43	12.64	12.64
under one year	1972	22.21	43.28	2130.39	32.62	45.27

. . . last page of table B.2 continued

Variable	Unweighted N	Unweighted Percent	Unweighted Cum.	Weighted N	Weighted Percent	Weighted Cum.
one year	3570	45.64	88.92	2786.08	42.67	87.93
more than one year	867	11.08	100.00	788.10	12.07	100.00
Total	7822	100.00		5924.00	100.00	
Estimate: exogenous, category						
missing	34	0.43	0.43	75.36	1.15	1.15
binary	914	11.68	12.12	1232.60	18.88	20.03
nominal	41	0.52	12.64	16.13	0.25	20.28
ordinal	750	9.59	22.23	483.20	7.40	27.68
metric	5116	65.41	87.64	4129.71	63.24	90.92
interval	967	12.36	100.00	593.00	9.08	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, in logs						
missing	455	5.82	5.82	306.76	4.70	4.70
yes	1571	20.08	25.90	1319.27	20.20	24.90
no	5796	74.10	100.00	4903.98	75.10	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, in differences						
missing	7589	97.02	97.02	6310.53	96.64	96.64
yes	227	2.90	99.92	204.20	3.13	99.77
no	6	0.08	100.00	15.28	0.23	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: exogenous, other transformation						
missing	433	5.54	5.54	288.89	4.42	4.42
yes	361	4.62	10.15	368.42	5.64	10.07
no	7028	89.85	100.00	5872.70	89.93	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: endogenous, index						
missing	7659	97.92	97.92	6424.22	98.38	98.38
additive, unweighted	49	0.63	98.54	49.75	0.76	99.14
additive, weighted	42	0.54	99.08	6.74	0.10	99.25
mean	58	0.74	99.82	30.30	0.46	99.71
multiplicative	9	0.12	99.94	13.58	0.21	99.92
miscellaneous	5	0.06	100.00	5.41	0.08	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: endogenous, number of reported crimes (absolute numbers)						
no	7230	92.43	92.43	5808.12	88.95	88.95
yes	592	7.57	100.00	721.88	11.05	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: endogenous, number of registered suspects						
no	7800	99.72	99.72	6492.61	99.43	99.43
yes	22	0.28	100.00	37.39	0.57	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: endogenous, number of incarcerated						
no	7816	99.92	99.92	6495.20	99.47	99.47
yes	6	0.08	100.00	34.80	0.53	100.00

. . . last page of table B.2 continued

Variable	Unweighted N	Unweighted Percent	Unweighted Cum.	Weighted N	Weighted Percent	Weighted Cum.
Total	7822	100.00		6530.00	100.00	
Estimate: endogenous, crime rate						
no	3745	47.88	47.88	3611.54	55.31	55.31
yes	4077	52.12	100.00	2918.46	44.69	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: endogenous, number of suspects						
no	7816	99.92	99.92	6524.18	99.91	99.91
yes	6	0.08	100.00	5.82	0.09	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: endogenous, number of convicted						
no	7807	99.81	99.81	6509.63	99.69	99.69
yes	15	0.19	100.00	20.37	0.31	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: endogenous, number of convicted to prison sentence						
no	7807	99.81	99.81	6516.42	99.79	99.79
yes	15	0.19	100.00	13.58	0.21	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: endogenous, self reported delinquency						
no	5921	75.70	75.70	5346.10	81.87	81.87
yes	1901	24.30	100.00	1183.90	18.13	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: endogenous, probability of future delinquency (surveyed is delinquent)						
no	7369	94.21	94.21	6263.63	95.92	95.92
yes	453	5.79	100.00	266.37	4.08	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: endogenous, probability of future delinquency (fictitious delinquent)						
no	7802	99.74	99.74	6509.27	99.68	99.68
yes	20	0.26	100.00	20.73	0.32	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: endogenous, probability of delinquency of fictitious offense (surveyed is delinquent)						
no	7761	99.22	99.22	6452.44	98.81	98.81
yes	61	0.78	100.00	77.56	1.19	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: endogenous, probability of delinquency of fictitious offense (surveyed is fictitious delinquent)						
no	7808	99.82	99.82	6480.10	99.24	99.24
yes	14	0.18	100.00	49.90	0.76	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: endogenous, other						
no	7550	96.52	96.52	5862.18	89.77	89.77
yes	272	3.48	100.00	667.82	10.23	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: endogenous, recidivism						
no	7788	99.57	99.57	6485.87	99.32	99.32

. . . last page of table B.2 continued

Variable	Unweighted N	Unweighted Percent	Unweighted Cum.	Weighted N	Weighted Percent	Weighted Cum.
yes	34	0.43	100.00	44.13	0.68	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: endogenous, accidents						
no	7632	97.57	97.57	6255.59	95.80	95.80
yes	190	2.43	100.00	274.41	4.20	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: endogenous, violating prescriptive limits						
no	7681	98.20	98.20	6343.36	97.14	97.14
yes	141	1.80	100.00	186.64	2.86	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: endogenous, relates to . . .						
missing	2040	26.08	26.08	1312.88	20.11	20.11
under one year	1704	21.78	47.86	1908.00	29.22	49.32
one year	3209	41.03	88.89	2555.30	39.13	88.46
more than one year	869	11.11	100.00	753.82	11.54	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: endogenous, self reported delinquency						
missing	5982	76.48	76.48	5421.42	83.02	83.02
under one year	329	4.21	80.68	288.55	4.42	87.44
one year	1134	14.50	95.18	535.66	8.20	95.65
since age of 14	32	0.41	95.59	39.04	0.60	96.24
lifelong	345	4.41	100.00	245.33	3.76	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: endogenous, future self reported delinquency						
missing	7339	93.83	93.83	6255.52	95.80	95.80
under one year	73	0.93	94.76	79.78	1.22	97.02
one year	77	0.98	95.74	57.29	0.88	97.90
more than one year	43	0.55	96.29	7.91	0.12	98.02
unlimited	290	3.71	100.00	129.51	1.98	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: endogenous and exogenous relate to the same offense						
missing	27	0.35	0.35	42.59	0.65	0.65
yes	7363	94.13	94.48	6075.90	93.05	93.70
no	432	5.52	100.00	411.52	6.30	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: crime category, crimes						
no	445	5.69	5.69	556.10	8.52	8.52
yes	7377	94.31	100.00	5973.90	91.48	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: crime category, misdemeanors						
no	7266	92.89	92.89	5910.26	90.51	90.51
yes	556	7.11	100.00	619.74	9.49	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: crime category, formal deviant behavior						
no	7653	97.84	97.84	6369.00	97.53	97.53

. . . last page of table B.2 continued

Variable	Unweighted N	Unweighted Percent	Unweighted Cum.	Weighted N	Weighted Percent	Weighted Cum.
yes	169	2.16	100.00	161.00	2.47	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: crime category, informal deviant behavior						
no	7684	98.24	98.24	6416.96	98.27	98.27
yes	138	1.76	100.00	113.04	1.73	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: crime category, violation of game-rules						
no	7781	99.48	99.48	6372.14	97.58	97.58
yes	41	0.52	100.00	157.86	2.42	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: crime category, other						
no	7762	99.23	99.23	6434.10	98.53	98.53
yes	60	0.77	100.00	95.90	1.47	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, homicide						
no	6121	78.25	78.25	5114.99	78.33	78.33
yes	1701	21.75	100.00	1415.01	21.67	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, manslaughter						
no	7752	99.11	99.11	6448.84	98.76	98.76
yes	70	0.89	100.00	81.16	1.24	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, assault						
no	6957	88.94	88.94	5869.42	89.88	89.88
yes	865	11.06	100.00	660.58	10.12	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, negligent assault						
no	7686	98.26	98.26	6410.74	98.17	98.17
yes	136	1.74	100.00	119.26	1.83	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, malicious mischief						
no	7659	97.92	97.92	6437.36	98.58	98.58
yes	163	2.08	100.00	92.64	1.42	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, burglary						
no	6970	89.11	89.11	5734.70	87.82	87.82
yes	852	10.89	100.00	795.30	12.18	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, robbery						
no	6852	87.60	87.60	5740.93	87.92	87.92
yes	970	12.40	100.00	789.07	12.08	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, larceny (Index I, general)						
no	6679	85.39	85.39	5709.09	87.43	87.43
yes	1143	14.61	100.00	820.91	12.57	100.00

. . . last page of table B.2 continued

Variable	Unweighted N	Percent	Cum.	Weighted N	Percent	Cum.
Total	7822	100.00		6530.00	100.00	
Estimate: offense, larceny (inferior)						
no	7521	96.15	96.15	6356.75	97.35	97.35
yes	301	3.85	100.00	173.25	2.65	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, larceny (severe)						
no	7519	96.13	96.13	6322.85	96.83	96.83
yes	303	3.87	100.00	207.15	3.17	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, drug possession (soft)						
no	7764	99.26	99.26	6485.11	99.31	99.31
yes	58	0.74	100.00	44.89	0.69	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, drug dealing (soft)						
no	7758	99.18	99.18	6488.48	99.36	99.36
yes	64	0.82	100.00	41.52	0.64	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, drug possession (hard)						
no	7804	99.77	99.77	6497.84	99.51	99.51
yes	18	0.23	100.00	32.16	0.49	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, drug dealing (hard)						
no	7782	99.49	99.49	6500.81	99.55	99.55
yes	40	0.51	100.00	29.19	0.45	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, drug related crime (general)						
no	7338	93.81	93.81	6241.83	95.59	95.59
yes	484	6.19	100.00	288.17	4.41	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, rape						
no	7320	93.58	93.58	6078.18	93.08	93.08
yes	502	6.42	100.00	451.82	6.92	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, sexual assault						
no	7751	99.09	99.09	6497.45	99.50	99.50
yes	71	0.91	100.00	32.55	0.50	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, other sexual related crimes						
no	7766	99.28	99.28	6479.50	99.23	99.23
yes	56	0.72	100.00	50.50	0.77	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, speeding						
no	7774	99.39	99.39	6457.53	98.89	98.89
yes	48	0.61	100.00	72.47	1.11	100.00
Total	7822	100.00		6530.00	100.00	

. . . last page of table B.2 continued

Variable	Unweighted N	Unweighted Percent	Unweighted Cum.	Weighted N	Weighted Percent	Weighted Cum.
Estimate: offense, driving without a licence						
no	7786	99.54	99.54	6527.27	99.96	99.96
yes	36	0.46	100.00	2.73	0.04	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, drunk driving						
no	7323	93.62	93.62	5742.83	87.95	87.95
yes	499	6.38	100.00	787.17	12.05	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, fare dodging						
no	7731	98.84	98.84	6506.06	99.63	99.63
yes	91	1.16	100.00	23.94	0.37	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, blackmailing						
no	7817	99.94	99.94	6513.87	99.75	99.75
yes	5	0.06	100.00	16.13	0.25	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, fraud						
no	7525	96.20	96.20	6273.26	96.07	96.07
yes	297	3.80	100.00	256.74	3.93	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, tax evasion						
no	7563	96.69	96.69	6056.33	92.75	92.75
yes	259	3.31	100.00	473.67	7.25	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, defalcation						
no	7806	99.80	99.80	6522.62	99.89	99.89
yes	16	0.20	100.00	7.38	0.11	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, embezzlement						
no	7800	99.72	99.72	6522.79	99.89	99.89
yes	22	0.28	100.00	7.21	0.11	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, smuggling						
no	7812	99.87	99.87	6528.07	99.97	99.97
yes	10	0.13	100.00	1.93	0.03	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, other crimes						
no	7312	93.48	93.48	6003.69	91.94	91.94
yes	510	6.52	100.00	526.31	8.06	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, crime rate (general)						
no	7501	95.90	95.90	6127.96	93.84	93.84
yes	321	4.10	100.00	402.04	6.16	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, other misdemeanors						

. . . last page of table B.2 continued

Variable	Unweighted N	Unweighted Percent	Unweighted Cum.	Weighted N	Weighted Percent	Weighted Cum.
no	7562	96.68	96.68	6323.89	96.84	96.84
yes	260	3.32	100.00	206.11	3.16	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, other						
no	7528	96.24	96.24	6069.87	92.95	92.95
yes	294	3.76	100.00	460.13	7.05	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, vehicle theft						
no	7174	91.72	91.72	5972.46	91.46	91.46
yes	648	8.28	100.00	557.54	8.54	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: offense, environmental crimes or violations of prescriptive limits						
no	7699	98.43	98.43	6378.70	97.68	97.68
yes	123	1.57	100.00	151.30	2.32	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, income inequality						
no	7076	90.46	90.46	6007.53	92.00	92.00
yes	746	9.54	100.00	522.47	8.00	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, time trend						
no	7530	96.27	96.27	6178.56	94.62	94.62
yes	292	3.73	100.00	351.44	5.38	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: characteristics						
missing	232	2.97	2.97	239.63	3.67	3.67
property offenses	2591	33.12	36.09	1959.87	30.01	33.68
violent offenses	1160	14.83	50.92	984.77	15.08	48.76
property and violent offenses	3839	49.08	100.00	3345.73	51.24	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: endogenous, category						
missing	27	0.35	0.35	60.09	0.92	0.92
binary	447	5.71	6.06	625.76	9.58	10.50
nominal	162	2.07	8.13	20.37	0.31	10.81
ordinal	276	3.53	11.66	299.77	4.59	15.41
metric	6597	84.34	96.00	5256.87	80.50	95.91
interval	313	4.00	100.00	267.15	4.09	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: endogenous, in logs						
missing	417	5.33	5.33	335.93	5.14	5.14
yes	2108	26.95	32.28	1750.43	26.81	31.95
no	5297	67.72	100.00	4443.64	68.05	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: endogenous, in differences						
missing	7588	97.01	97.01	6313.09	96.68	96.68
yes	234	2.99	100.00	216.91	3.32	100.00

. . . last page of table B.2 continued

Variable	Unweighted N	Unweighted Percent	Unweighted Cum.	Weighted N	Weighted Percent	Weighted Cum.
Total	7822	100.00		6530.00	100.00	
Estimate: endogenous, other transformation						
missing	444	5.68	5.68	327.85	5.02	5.02
yes	524	6.70	12.38	509.04	7.80	12.82
no	6854	87.62	100.00	5693.12	87.18	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: relation between endogenous and exogenous						
missing	42	0.54	0.54	70.29	1.08	1.08
same time	5702	72.90	73.43	4701.48	72.00	73.07
exogenous before endogenous	1967	25.15	98.58	1674.51	25.64	98.72
exogenous after endogenous	111	1.42	100.00	83.72	1.28	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, age						
no	6389	81.68	81.68	5210.87	79.80	79.80
yes	1433	18.32	100.00	1319.13	20.20	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, sex						
no	6828	87.29	87.29	5593.92	85.66	85.66
yes	994	12.71	100.00	936.08	14.34	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, marital status						
no	7507	95.97	95.97	6162.82	94.38	94.38
yes	315	4.03	100.00	367.18	5.62	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, nationality						
no	7361	94.11	94.11	6317.97	96.75	96.75
yes	461	5.89	100.00	212.03	3.25	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, education						
no	7085	90.58	90.58	5960.68	91.28	91.28
yes	737	9.42	100.00	569.32	8.72	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, income						
no	5545	70.89	70.89	4609.52	70.59	70.59
yes	2277	29.11	100.00	1920.48	29.41	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, profession						
no	7805	99.78	99.78	6505.39	99.62	99.62
yes	17	0.22	100.00	24.61	0.38	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, unemployment						
no	5378	68.75	68.75	4787.35	73.31	73.31
yes	2444	31.25	100.00	1742.65	26.69	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, social integration						

. . . last page of table B.2 continued

Variable	Unweighted N	Unweighted Percent	Unweighted Cum.	Weighted N	Weighted Percent	Weighted Cum.
no	7655	97.86	97.86	6432.71	98.51	98.51
yes	167	2.14	100.00	97.29	1.49	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, religion						
no	7636	97.62	97.62	6418.28	98.29	98.29
yes	186	2.38	100.00	111.72	1.71	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, social class						
no	7735	98.89	98.89	6477.80	99.20	99.20
yes	87	1.11	100.00	52.20	0.80	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, drug usage						
no	7743	98.99	98.99	6445.90	98.71	98.71
yes	79	1.01	100.00	84.10	1.29	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, previous convictions						
no	7642	97.70	97.70	6424.22	98.38	98.38
yes	180	2.30	100.00	105.78	1.62	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, previous incarceration						
no	7797	99.68	99.68	6472.29	99.12	99.12
yes	25	0.32	100.00	57.71	0.88	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, acceptance of norms						
no	7411	94.75	94.75	6388.11	97.83	97.83
yes	411	5.25	100.00	141.89	2.17	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, morality						
no	7601	97.17	97.17	6424.12	98.38	98.38
yes	221	2.83	100.00	105.88	1.62	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, personal characteristics						
no	7662	97.95	97.95	6337.20	97.05	97.05
yes	160	2.05	100.00	192.80	2.95	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, importance of goal						
no	7822	100.00	100.00	6530.00	100.00	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, possibility of legal success						
no	7814	99.90	99.90	6524.91	99.92	99.92
yes	8	0.10	100.00	5.09	0.08	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, utility from offenses						
no	7746	99.03	99.03	6507.93	99.66	99.66
yes	76	0.97	100.00	22.07	0.34	100.00

...last page of table B.2 continued

	Unweighted			Weighted		
Variable	N	Percent	Cum.	N	Percent	Cum.
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, utility from legal work						
no	7822	100.00	100.00	6530.00	100.00	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, costs of legal success						
no	7822	100.00	100.00	6530.00	100.00	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, costs from illegal success						
no	7814	99.90	99.90	6524.91	99.92	99.92
yes	8	0.10	100.00	5.09	0.08	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, fixed effects (spatial)						
no	6840	87.45	87.45	5879.13	90.03	90.03
yes	982	12.55	100.00	650.87	9.97	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, fixed effects (time)						
no	6726	85.99	85.99	5747.60	88.02	88.02
yes	1096	14.01	100.00	782.40	11.98	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, random effects						
no	7780	99.46	99.46	6464.05	98.99	98.99
yes	42	0.54	100.00	65.95	1.01	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, other						
no	3087	39.47	39.47	2339.57	35.83	35.83
yes	4735	60.53	100.00	4190.43	64.17	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, youths						
no	6710	85.78	85.78	5609.93	85.91	85.91
yes	1112	14.22	100.00	920.07	14.09	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, race						
no	5953	76.11	76.11	5062.33	77.52	77.52
yes	1869	23.89	100.00	1467.67	22.48	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, poverty or welfare						
no	7320	93.58	93.58	6113.61	93.62	93.62
yes	502	6.42	100.00	416.39	6.38	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, urbanity						
no	6970	89.11	89.11	5996.94	91.84	91.84
yes	852	10.89	100.00	533.06	8.16	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, GDP						
no	7714	98.62	98.62	6441.31	98.64	98.64

. . . last page of table B.2 continued

Variable	Unweighted N	Unweighted Percent	Unweighted Cum.	Weighted N	Weighted Percent	Weighted Cum.
yes	108	1.38	100.00	88.69	1.36	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, population (-growth)						
no	6755	86.36	86.36	5782.27	88.55	88.55
yes	1067	13.64	100.00	747.73	11.45	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, alcohol (consumption)						
no	7724	98.75	98.75	6421.01	98.33	98.33
yes	98	1.25	100.00	108.99	1.67	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, labor force						
no	7670	98.06	98.06	6365.06	97.47	97.47
yes	152	1.94	100.00	164.94	2.53	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, property value						
no	7753	99.12	99.12	6407.45	98.12	98.12
yes	69	0.88	100.00	122.55	1.88	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, consumption						
no	7720	98.70	98.70	6400.18	98.01	98.01
yes	102	1.30	100.00	129.82	1.99	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, risk propensity						
no	7781	99.48	99.48	6476.18	99.18	99.18
yes	41	0.52	100.00	53.82	0.82	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: covariate, miles driven						
no	7728	98.80	98.80	6444.68	98.69	98.69
yes	94	1.20	100.00	85.32	1.31	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: linear model						
missing	328	4.19	4.19	282.36	4.32	4.32
yes	6720	85.91	90.10	5264.25	80.62	84.94
no	774	9.90	100.00	983.39	15.06	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: additive model						
missing	355	4.54	4.54	301.95	4.62	4.62
yes	7388	94.45	98.99	6121.57	93.75	98.37
no	79	1.01	100.00	106.48	1.63	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: error corrections implemented						
yes	1037	13.26	13.26	837.09	12.82	12.82
no	6785	86.74	100.00	5692.91	87.18	100.00
Total	7822	100.00		6530.00	100.00	

. . . last page of table B.2 continued

Variable	Unweighted N	Percent	Cum.	Weighted N	Percent	Cum.
Estimate: weighted model						
yes	697	8.91	8.91	445.06	6.82	6.82
no	7125	91.09	100.00	6084.94	93.18	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: correction for simultaneity						
missing	4359	55.73	55.73	3120.28	47.78	47.78
by methodology	638	8.16	63.88	594.60	9.11	56.89
with variables	645	8.25	72.13	593.49	9.09	65.98
by methodology and variables	173	2.21	74.34	74.33	1.14	67.12
none	2007	25.66	100.00	2147.30	32.88	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: bivariate method						
missing	5352	68.42	68.42	4936.38	75.60	75.60
chi^2	36	0.46	68.88	97.09	1.49	77.08
contingency coefficient	7	0.09	68.97	5.09	0.08	77.16
phi	48	0.61	69.59	15.01	0.23	77.39
binomial test	6	0.08	69.66	10.18	0.16	77.55
Wilcoxon test	2	0.03	69.69	10.18	0.16	77.70
other non-parametric test	5	0.07	69.75	5.82	0.09	77.79
Spearman's ρ	149	1.90	71.66	29.71	0.45	78.25
Kendall's τ	197	2.52	74.18	55.83	0.85	79.10
γ	214	2.74	76.91	148.00	2.27	81.37
t-test for independent samples	6	0.08	76.99	35.22	0.54	81.91
t-test for dependant samples	9	0.12	77.10	36.83	0.56	82.47
ANOVA	36	0.46	77.56	139.87	2.14	84.61
differences	104	1.33	78.89	258.23	3.95	88.57
point biserial correlation	85	1.09	79.98	24.75	0.38	88.95
Pearson correlation	1462	18.69	98.67	617.10	9.45	98.40
regression	70	0.89	99.57	41.87	0.64	99.04
other	34	0.43	100.00	62.83	0.96	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: multivariate method						
missing	2473	31.62	31.62	1593.2	24.40	24.40
VAR	46	0.59	32.20	41.59	0.64	25.04
ANOVA, ANCOVA	7	0.09	32.29	25.46	0.39	25.43
Logit/Probit (standardized)	120	1.53	33.83	107.08	1.64	27.07
Logit/Probit (unstandardized)	286	3.66	37.48	393.65	6.03	33.10
partial correlation	167	2.14	39.62	41.72	0.64	33.74
Poisson-regression	61	0.78	40.40	68.74	1.05	34.79
OLS (standardized)	1225	15.66	56.06	811.71	12.43	47.22
OLS (unstandardized)	1787	22.85	78.91	1582.90	24.24	71.46
2SLS (standardized)	43	0.55	79.46	53.07	0.81	72.27
2SLS (unstandardized)	735	9.40	88.85	638.88	9.78	82.06
Pathanalysis (standardized)	53	0.68	89.53	87.17	1.33	83.39
Pathanalysis (unstandardized)	9	0.12	89.64	10.18	0.16	83.55

. . . last page of table B.2 continued

Variable	Unweighted N	Unweighted Percent	Unweighted Cum.	Weighted N	Weighted Percent	Weighted Cum.
GLS	70	0.89	90.54	109.26	1.67	85.22
COX	15	0.19	90.73	20.37	0.31	85.53
GMM	35	0.45	91.18	50.68	0.78	86.31
ARIMA	132	1.69	92.87	275.93	4.23	90.54
TOBIT	180	2.30	95.17	215.77	3.30	93.84
other ML	322	4.12	99.28	331.37	5.07	98.92
other	56	0.72	100.00	70.84	1.08	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: sign supports deterrence hypothesis						
missing	181	2.31	2.31	0.00	0.00	0.00
yes	5459	69.79	72.10	4820.64	73.82	73.82
no	2182	27.90	100.00	1709.36	26.18	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: p-value in original study						
no	932	11.92	11.92	357.25	5.47	5.47
yes	6890	88.08	100.00	6172.75	94.53	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: p-value						
missing	765	9.78	9.78	0.00	0.00	0.00
99.9%, support	622	7.95	17.73	671.54	10.28	10.28
99%, support	760	9.72	27.45	835.06	12.79	23.07
95%, support	1111	14.20	41.65	1215.75	18.62	41.69
90%, support	268	3.43	45.08	278.63	4.27	45.96
not significant, support	2307	29.49	74.57	1819.66	27.87	73.82
not significant, no support	1464	18.72	93.29	1141.63	17.48	91.31
90%, no support	57	0.73	94.02	52.32	0.80	92.11
95%, no support	256	3.27	97.29	282.24	4.32	96.43
99%, no support	119	1.52	98.81	137.51	2.11	98.54
99.9%, no support	93	1.19	100.00	95.66	1.46	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: t-value in original study						
no	5574	71.26	71.26	3868.42	59.24	59.24
yes	2248	28.74	100.00	2661.58	40.76	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: F-value in study						
no	7653	97.84	97.84	6315.27	96.71	96.71
yes	169	2.16	100.00	214.73	3.29	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: χ^2-value in study						
no	7772	99.36	99.36	6386.57	97.80	97.80
yes	50	0.64	100.00	143.43	2.20	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: other-value in study						
no	7758	99.18	99.18	6380.63	97.71	97.71
yes	64	0.82	100.00	149.37	2.29	100.00

. . . last page of table B.2 continued

Variable	Unweighted N	Percent	Cum.	Weighted N	Percent	Cum.
Total	7822	100.00		6530.00	100.00	
Estimate: violent crime						
no	4947	63.24	63.24	4267.44	65.35	65.35
yes	2875	36.76	100.00	2262.56	34.65	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: property crime						
no	3850	49.22	49.22	3308.39	50.66	50.66
yes	3972	50.78	100.00	3221.61	49.34	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: other crime						
no	5774	73.82	73.82	4261.11	65.25	65.25
yes	2048	26.18	100.00	2268.89	34.75	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: dummy for first estimate						
no	7122	91.05	91.05	4584.33	70.20	70.20
yes	700	8.95	100.00	1945.67	29.80	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: dummy for randomly chosen or unique estimate						
no	4527	57.88	57.88	2114.23	32.38	32.38
yes	3295	42.12	100.00	4415.77	67.62	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: dummy for favored and not unique estimate						
no	6142	78.52	78.52	5154.52	78.94	78.94
yes	1680	21.48	100.00	1375.48	21.06	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: negative and significant (by p-value)						
missing	374	4.78	4.78	0.00	0.00	0.00
no	4930	63.03	63.03	3788.77	58.02	58.02
yes	2518	32.19	100.00	2741.23	41.98	100.00
Total	7822	100.00		6530.00	100.00	
Estimate: negative and significant (by normalized t-value)						
missing	765	9.78	9.78	0.00	0.00	0.00
no	4538	58.02	67.80	3809.92	58.34	58.34
yes	2519	32.20	100.00	2720.08	41.66	100.00
Total	7822	100.00		6530.00	100.00	

Columns 2-4 are not weighted and refer to all estimates. Columns 5-7 are weighted and relate to all valid estimates.

Tables B.3 and B.4 present the metric variables (neglecting missing values).

Table B.3: Descriptive statistics of the metric study-variables

	Min	Mean	Median	Max	sd	Weights	N
Study: publication, volume							
	1	37.95	31	455	33.08	604	604

...last page of table B.3 continued

	Min	Mean	Median	Max	sd	Weights	N
	1	38.28	31	455	33.38	553.30	5126
Study: publication, number							
	1	4.91	3	437	27.49	584	584
	1	5.11	3	437	28.69	534.84	5008
Study: publication, year							
	1952	1989.78	1990	2006	10.60	700	700
	1952	1990.01	1990	2006	10.55	641.18	6530
Study: publication, page start							
	1	303.47	241	1869	294.06	699	699
	1	302.03	237	1869	294.66	640.18	6526
Study: publication, page end							
	7	326.79	264	1879	289.25	700	700
	7	325.53	257	1879	289.91	641.18	6530
Study: publication, number of pages							
	2	24.75	17	439	38.05	700	700
	2	24.97	17	439	38.45	641.18	6530
Study: measuring points							
	1	36.33	3	4197	212.18	682	682
	1	37.96	3	4197	221.12	624.30	6408
Study: year of first measure							
	1830	1971.35	1975	2002	19.93	593	593
	1830	1971.74	1975	2002	19.88	540.92	5273
Study: year of last measure							
	1892	1983.48	1986	2004	14.08	428	428
	1892	1983.70	1987	2004	14.20	391.98	3957
Study: time span in months							
	1	185.00	120	1728	195.70	442	442
	1	184.85	120	1728	194.07	404.92	4347
Study: number of studied populations							
	1	1.07	1	7	0.39	700	700
	1	1.06	1	7	0.36	641.18	6530
Study: size, first population							
	1	17677.39	229	2167700	169012.90	168	168
	1	8429.98	212	2167700	101642.08	159.76	2449
Study: size, second population							
	44	649.14	262	3145	1106.10	7	7
	44	715.12	262	3145	1104.80	5.91	178
Study: size, first sample							
	1	4344.02	245	199358	20280.04	178	178
	1	4521.18	222	199358	21003.51	164.03	1398
Study: size, second sample							
	13	14685.78	1725	101636	33433.50	9	9
	13	16114.44	1725	101636	33506.94	7.82	127
Study: size, first realized sample							
	1	1763.98	133	135931	9023.28	398	398
	1	1736.17	130	135931	9216.19	370.27	3842

. . . last page of table B.3 continued

	Min	Mean	Median	Max	sd	Weights	N
Study: size, second realized sample							
	13	1679.61	293	16193	3755.34	18	18
	13	1490.56	262	16193	3437.58	15.53	331
Study: rate of return, first sample							
	13	76.80	81	100	22.53	90	90
	13	75.98	80	100	22.78	83.14	764
Study: rate of return, second sample							
	49	79.00	85	100	22.39	7	7
	49	76.6	85	100	21.19	5.95	35
Study: minimum age, first population							
	10	16.44	17	24	2.81	61	61
	10	16.41	17	24	2.85	57.60	624
Study: minimum age, second population							
	11	14.50	15	18	3.11	4	4
	11	14.51	16	18	2.71	3.98	144
Study: maximum age, first population							
	16	35.21	24	99	22.58	39	39
	16	35.16	24	99	22.60	36.82	367
Study: maximum age, second population							
	18	21.00	21	24	4.24	2	2
	18	21.04	24	24	3.01	1.98	136
Study: mean age, first population							
	15	26.29	22	76	12.57	34	34
	15	26.65	22	76	12.72	31.85	289
Study: mean age, second population							
	17	22.00	22	27	7.07	2	2
	17	22.00	22	27	5.35	2.00	8
Study: female fraction, first sample							
	0	33.79	46	84	24.17	87	87
	0	32.81	46	84	23.90	82.00	952
Study: female fraction, second sample							
	0	34.30	48.39	58	26.82	6	6
	0	34.23	46.78	58	24.59	5.98	149
Study: number of estimates, bivariate							
	1	12.24	5	155	21.91	238	238
	1	11.08	4	155	19.04	208.01	2889
Study: number of estimates, multivariate							
	1	21.72	8	764	52.02	570	570
	1	21.70	8	764	53.27	530.79	5782

Each first row refers to all studies (maximum of 700) and is unweighted. The second row is weighted (each study is weighted equally, the maximum cumulative weight is 641.18) and relates to all valid estimates (maximum of 6530).
Missing values are not considered.

Table B.4: Descriptive statistics of the metric estimate-variables

	Min	Mean	Median	Max	sd	Weights	N
Estimate: number of estimates							
	1.00	21.19	10.00	210.00	28.90	7822	7822
	1.00	6.01	3.00	210.00	10.77	641.18	6530
Estimate: exogenous, number of index items							
	1.00	1.07	1.00	21.00	0.57	7818	7818
	1.00	1.06	1.00	9.00	0.44	641.18	6530
Estimate: exogenous, first year of observation							
	1830.00	1968.42	1973.00	2002.00	21.78	6130	6130
	1830.00	1972.14	1975.00	2002.00	19.26	517.54	5027
Estimate: exogenous, last year of observation							
	1892.00	1979.87	1983.00	2004.00	16.66	5454	5454
	1892.00	1982.91	1986.00	2004.00	14.44	453.09	4495
Estimate: exogenous, number of categories							
	2.00	27.86	5.00	100.00	41.14	2680	2680
	2.00	16.54	2.00	100.00	33.07	226.28	2442
Estimate: endogenous, number of index items							
	1.00	1.06	1.00	14.00	0.51	7822	7822
	1.00	1.06	1.00	14.00	0.64	641.18	6530
Estimate: endogenous, first year of observation							
	1830.00	1968.74	1973.00	2002.00	21.39	5759	5759
	1830.00	1972.21	1975.00	2002.00	19.11	503.30	4656
Estimate: endogenous, last year of observation							
	1892.00	1980.58	1985.00	2004.00	16.81	5025	5025
	1892.00	1983.26	1986.00	2004.00	14.46	435.38	4086
Estimate: endogenous, number of categories							
	2.00	12.25	3.00	200.00	29.13	1170	1170
	0.00	2.50	0.00	200.00	15.44	641.18	6530
Estimate: number of used covariates							
	1.00	7.55	7.00	92.00	6.05	5297	5297
	1.00	7.14	6.00	92.00	6.20	481.76	4470
Estimate: sd of estimate							
	0.00	4.78	0.12	861.40	49.55	1177	1177
	0.00	5.54	0.15	861.40	49.76	103.61	913
Estimate: effect value							
	−10122.00	2.95	−0.07	40001.04	527.71	7374	7374
	−10122.00	24.91	−0.08	40001.04	1213.91	574.45	6288
Estimate: degrees of freedom of effect value							
	305.00	305.00	305.00	305.00	0.00	1	1
	305.00	305.00	305.00	305.00	0.00	1.00	1
Estimate: t-value, not normalized							
	−582.00	−1.45	−0.97	21.54	7.99	7057	7057
	−582.00	−1.51	−1.37	20.93	4.51	641.18	6530
Estimate: degrees of freedom of t-value							
	4.00	158.00	10.00	1102.00	327.38	59	59

. . . last page of table B.4 continued

	Min	Mean	Median	Max	sd	Weights	N
	4.00	246.52	24.00	1102.00	389.83	10.27	57
Estimate: value of F-value							
	−19.55	22.34	2.14	1745.73	140.14	168	168
	−19.55	67.16	4.35	1745.73	275.41	20.58	163
Estimate: degrees of freedom of F-value							
	1.00	13.09	1.00	132.00	39.44	11	11
	1.00	13.81	1.00	132.00	40.58	5.17	11
Estimate: value of χ^2-value							
	0.01	9838.85	9.90	240498.00	48074.47	49	49
	0.01	17205.79	9.90	240498.00	62543.34	14.01	49
Estimate: degrees of freedom of χ^2-value							
	1.00	5.41	7.00	11.00	2.79	27	27
	1.00	4.04	2.00	11.00	3.60	8.17	27
Estimate: value of other-value							
	−1.38	3.79	0.22	83.00	14.22	64	64
	−1.38	2.60	0.34	64.00	10.48	14.67	42
Estimate: degrees of freedom of other-value							
	1.00	1.00	1.00	1.00	0.00	1	1
	1.00	1.00	1.00	1.00	0.00	0.50	1
Estimate: R^2 of model							
	−0.06	0.62	0.71	1.00	0.30	3059	3059
	−0.06	0.61	0.68	1.00	0.30	253.22	2642
Estimate: F-value of model							
	0.69	57.20	24.30	797.50	92.22	642	642
	0.69	58.55	23.80	797.50	104.25	65.89	557
Estimate: goodness of fit value							
	0.01	27.08	1.73	625.45	91.10	48	48
	0.01	20.26	1.00	625.45	72.83	6.00	41
Estimate: degrees of freedom of model							
	4.00	223.94	36.00	3439.00	531.20	383	383
	4.00	268.27	36.00	3439.00	577.52	48.38	341
Estimate: other model value							
	−21284.50	−212.52	1.71	33532.00	2707.48	692	692
	−21284.50	−221.79	1.76	33532.00	3584.15	76.41	614
Estimate: sample size							
	5.00	1518.60	232.00	206035.00	6531.11	5972	5972
	5.00	2203.89	213.00	206035.00	11256.29	522.20	5050
Estimate: square root of sample size for positive values							
	3.00	21.10	13.13	223.62	28.70	1708	1708
	3.00	25.31	15.07	223.62	32.55	132.23	1404
Estimate: square root of sample size for negative values							
	2.24	23.69	15.84	453.91	32.53	4264	4264
	2.24	25.74	14.42	453.91	41.49	389.97	3646
Estimate: weights, weights all estimates equally							
	0.00	1.95	1.00	44.17	3.45	7822	7822
	1.00	2.48	1.00	44.17	3.83	641.18	6530

. . . last page of table B.4 continued

	Min	Mean	Median	Max	sd	Weights	N
Estimate: weights, weights all studies equally							
	0.00	0.09	0.04	1.00	0.14	7822	7822
	0.00	0.31	0.20	1.00	0.30	641.18	6530
Estimate: transformed p-values							
	0.00	0.82	1.00	1.00	0.39	7057	7057
	0.00	0.78	1.00	1.00	0.41	641.18	6530
Estimate: normalized t-value							
	−64.81	−1.27	−0.96	20.37	3.13	7057	7057
	−64.81	−1.40	−1.37	19.05	3.07	641.18	6530

Each first row refers to all observations of all studies and is unweighted (maximum of 7822; the weights are equal to the number of observations). The second row is weighted (maximum cumulative weight is 641.18) and relates to all valid estimates (maximum of 6530).
Missing values are not considered.

B.2 Included Studies

Table B.5 shows all 700 studies of our meta analysis. The first column contains the names of up to two authors. The crime type (CT) states whether violent, property or other crimes are studied (either yes or no). The overall opinion of the authors (O) ranges from -2, full support, to +2, no support; see subsection 3.3.3. The fourth column contains the abbreviation of the user who read the study and recorded its information; all but "tr" were part of the team in Heidelberg.

Table B.5: Studies included in the meta-data base

Authors	CT	O	U	Source
Schuessler, K.	nny	−1	aw	Schuessler (1952)
Michaels, R.	ynn	+1	tr	Michaels (1960)
Rettig, S.; Rawson, H.	nyn	−1	ah	Rettig and Rawson (1963)
Sinha, J.; Wherry, R.	ynn	−1	ah	Sinha and Wherry (1965)
Clarke, R.	ynn	−2	tr	Clarke (1966)
Sharkansky, I.	nyy	+2	ah	Sharkansky (1967)
Sinha, J.	ynn	−1	ah	Sinha (1967)
Campbell, D.; Ross, L.	ynn	−1	aw	Campbell and Ross (1968)
Cardarelli, A.	nny	+2	ah	Cardarelli (1968)
Glass, G.	ynn	−2	tr	Glass (1968)
Gray, L.; Martin, D.	nny	−1	aw	Gray and Martin (1969)
Hill, J.; Kochendorfer, R.	nyn	−1	aw	Hill and Kochendorfer (1969)
Horai, J.; Tedeschi, J.	ynn	−1	kr	Horai and Tedeschi (1969)
Jensen, G.	yyy	−1	aw	Jensen (1969)
Tittle, C.	nyy	−1	aw	Tittle (1969)
Chiricos, T.; Waldo, G.	nyy	0	kr	Chiricos and Waldo (1970)

...last page of table B.5 continued

Authors	CT	O	U	Source
Gahagan, J. et al.	ynn	+1	ah	Gahagan et al. (1970)
Ross, L. et al.	ynn	−1	aw	Ross et al. (1970)
Salem, R.; Bowers, W.	yyn	+2	tr	Salem and Bowers (1970)
Bailey, W.	nyy	0	aw	Bailey (1971)
Bean, F.; Cushing, R.	nny	−1	aw	Bean and Cushing (1971)
Logan, C.	yyy	+1	aw	Logan (1973)
Morris, D.; Tweeten, L.	nyy	0	tr	Morris and Tweeten (1971)
Press, S.	yyy	−2	tr	Press (1971)
Allison, J.	yyy	0	tr	Allison (1972)
Bowers, W.; Salem, R.	ynn	−1	ah	Bowers and Salem (1972)
Cho, Y.	nyy	+2	ah	Cho (1972)
Ehrlich, I.	nyy	−2	tr	Ehrlich (1972)
Logan, C.	nyy	−1	aw	Logan (1972)
Snyder, D.; Tilly, C.	ynn	+2	ah	Snyder and Tilly (1972)
Teevan, J.	yyy	+2	ah	Teevan (1972)
Waldo, G.; Chiricos, T.	yyn	+1	aw	Waldo and Chiricos (1972)
Antunes, G.; Hunt, A.	nyy	0	aw	Antunes and Hunt (1973a)
Antunes, G.; Hunt, A.	nyy	+1	aw	Antunes and Hunt (1973b)
Carr-Hill, R.; Stern, N.	yyy	−2	tr	Carr-Hill and Stern (1973)
Ehrlich, I.	nyy	−2	tr	Ehrlich (1973)
Erickson, M.; Gibbs, J.	nny	−1	aw	Erickson and Gibbs (1973)
Greenwood, M.; Wadycki, W.	nyy	0	tr	Greenwood and Wadycki (1973)
Jayewardene, C.	nny	+1	ah	Jayewardene (1973)
Jones, T.	nyy	+2	tr	Jones (1973)
Robertson, L. et al.	ynn	+2	tr	Robertson et al. (1973)
Tittle, C.; Rowe, A.	ynn	−1	ah	Tittle and Rowe (1973)
Bailey, W.	nny	+1	aw	Bailey (1974)
Bailey, W. et al.	nyy	−1	kr	Bailey et al. (1974)
Buikhuisen, W.	ynn	−2	tr	Buikhuisen (1974)
Chaiken, J. et al.	nyn	−2	tr	Chaiken et al. (1974)
Glaser, D.; Zeigler, M.	nny	+2	aw	Glaser and Zeigler (1974)
Heisler, G.	nyn	0	aw	Heisler (1974)
McPheters, L.; Stronge, W.	yyy	−2	tr	McPheters and Stronge (1974)
Swimmer, E.	nyy	−2	tr	Swimmer (1974a)
Swimmer, E.	nyy	+1	aw	Swimmer (1974b)
Tittle, C.; Rowe, A.	yyy	−1	mw	Tittle and Rowe (1974)
Wellford, C.	yyy	−1	aw	Wellford (1974)
Bacon, P.; O'Donoghue, M.	ynn	−2	tr	Bacon and O'Donohue (1975)
Bailey, W.	nny	−1	ah	Bailey (1975)
Bowers, W.; Pierce, G.	nny	+2	ah	Bowers and Pierce (1975)
Burkett, S.; Jensen, E.	ynn	+1	tr	Burkett and Jensen (1975)
Cloninger, D.	nyy	−1	tr	Cloninger (1975)
Danziger, S.; Wheeler, D.	nyy	−1	tr	Danziger and Wheeler (1975)
Ehrlich, I.	nny	−2	tr	Ehrlich (1975a)
Erickson, M.; Gibbs, J.	nyy	0	aw	Erickson and Gibbs (1975)
Greenwood, M.; Wadycki, W.	nyy	0	tr	Greenwood and Wadycki (1975)

. . . last page of table B.5 continued

Authors	CT	O	U	Source
Kau, J.; Rubin, P.	nyy	−1	tr	Kau and Rubin (1975)
Logan, C.	nyy	−1	aw	Logan (1975)
Minor, W.	yyn	+1	aw	Minor (1975)
Passell, P.	nny	−1	aw	Passell (1975)
Phillips, L.; Votey, H.	nyy	−2	ah	Phillips and Votey (1975)
Pogue, T.	yyy	0	tr	Pogue (1975)
Ross, L.	ynn	−1	ah	Ross (1975)
Williams, A.; Robertson, L.	ynn	+2	aw	Williams and Robertson (1975)
Avio, K.; Clark, C.	nyn	−1	tr	Avio and Clark (1976)
Bailey, W.	nny	+2	aw	Bailey (1976)
Bailey, W.; Lott, Ruth P.	yyn	+1	mw	Bailey and Lott (1976)
Chambers, L. et al.	ynn	−1	tr	Chambers et al. (1976)
Chapman, J.	nyy	−1	tr	Chapman (1976)
Erickson, M.; Gibbs, J.	nyy	−1	ah	Erickson and Gibbs (1976)
Erickson, P.	ynn	+2	aw	Erickson (1976)
Forst, B.	nyy	+1	aw	Forst (1976)
Grasmick, H.; Milligan, H.	ynn	−1	aw	Grasmick and Milligan (1976)
Land, K.; Felson, M.	nyy	0	aw	Land and Felson (1976)
Mathieson, D.; Passell, P.	nyy	−1	aw	Mathieson and Passell (1976)
Silberman, M.	yyy	0	kr	Silberman (1976)
Spicer, M.; Lundstedt, S.	nyn	+1	tr	Spicer and Lundstedt (1976)
Teevan, J.	nyn	+1	aw	Teevan (1976a)
Teevan, J.	yyn	+1	aw	Teevan (1976b)
Teevan, J.	yyn	+1	aw	Teevan (1976c)
Tittle, C.	yyy	−1	aw	Tittle (1974)
Upper, J.; White, J.	nyn	+2	ah	Upper and White (1976)
Yunker, J.	nny	−2	tr	Yunker (1976)
Zador, P.	ynn	+2	tr	Zador (1976)
Alcorn, D.	nyy	+1	aw	Alcorn (1977)
Anderson, L.	yyn	0	aw	Anderson (1978)
Anderson, L. et al.	ynn	0	aw	Anderson et al. (1977)
Bailey, W.	nny	−1	aw	Bailey (1977)
Blumstein, A.; Nagin, D.	ynn	+1	ah	Blumstein and Nagin (1977)
Cloninger, D.	nny	−2	tr	Cloninger (1977)
Ehrlich, I.	nyy	−2	tr	Ehrlich (1977a)
Ehrlich, I.	nny	−2	tr	Ehrlich (1977b)
Erickson, M. et al.	nyy	+1	tr	Erickson et al. (1977)
Forst, B.	nny	+2	ah	Forst (1977)
Fox, J.	nny	+2	tr	Fox (1977)
Geerken, M.; Gove, W.	nyy	−1	aw	Geerken and Gove (1977)
Grasmick, H.; Appleton, L.	ynn	−1	aw	Grasmick and Appleton (1977)
Sesnowitz, M.; McKee, D.	nny	+2	tr	Sesnowitz and McKee (1977)
Mehay, S.	nyy	−1	tr	Mehay (1977)
Meier, R.; Johnson, W.	ynn	+1	aw	Meier and Johnson (1977)
Nagel, W.	yyy	+2	ah	Nagel (1977)
Passell, P.; Taylor, J.	nny	+1	tr	Passell and Taylor (1977)

. . . last page of table B.5 continued

Authors	CT	O	U	Source
Ross, L.	ynn	−1	aw	Ross (1977)
Silverman, L.; Spruill, N.	nyy	−1	tr	Silverman and Spruill (1977)
Thaler, R.	nyn	−1	tr	Thaler (1977)
Victor, M.	nyy	+2	aw	Victor (1977)
Witte, A.; Schmidt, P.	yyy	0	tr	Witte and Schmidt (1977)
Avio, K.; Clark, C.	nyn	−1	tr	Avio and Clark (1978)
Bailey, W.	nny	+2	aw	Bailey (1978a)
Bailey, W.	nyy	0	ah	Bailey (1978b)
Bailey, W.	nny	+1	aw	Bailey (1978c)
Black, T.; Orsagh, T.	nny	−1	aw	Black and Orsagh (1978)
Brown, D.	nyy	+1	ah	Brown (1978)
Cohen, L.	ynn	+1	aw	Cohen (1978)
Erickson, M.; Gibbs, J.	yyy	−1	aw	Erickson and Gibbs (1978)
Friedland, N. et al.	nyn	−1	tr	Friedland et al. (1978)
Hakim, S. et al.	nyn	−2	tr	Hakim et al. (1978)
Holtmann, A.; Yap, L.	nyn	−2	tr	Holtmann and Yap (1978)
Hurst, P.	ynn	−1	ah	Hurst (1978)
Jensen, G. et al.	yyn	0	aw	Jensen et al. (1978)
Klemke, L.	nyn	−1	aw	Klemke (1978)
Landes, W.	ynn	−2	tr	Landes (1978)
Levy, P. et al.	ynn	−1	aw	Levy et al. (1978)
Mason, R.; Calvin, L.	nyn	−1	aw	Mason and Calvin (1978)
Mathur, V.	nyy	−2	tr	Mathur (1978)
McPheters, L.	yyy	+2	ah	McPheters (1978)
Norström, T.	ynn	−1	ah	Norström (1978)
Pontell, H.	nyy	+1	ah	Pontell (1978)
Vandaele, W.	nyn	−2	tr	Vandaele (1978a)
Vandaele, W.	nyy	−2	tr	Vandaele (1978b)
Votey, H.	ynn	−2	tr	Votey (1978)
Wilson, J.; Boland, B.	nyn	−2	aw	Wilson and Boland (1978)
Wolpin, K.	nny	−2	tr	Wolpin (1978a)
Wolpin, K.	yyy	+1	tr	Wolpin (1978b)
Akers, R. et al.	ynn	+2	aw	Akers et al. (1979)
Archambeault, W.	ynn	−2	aw	Archambeault (1979)
Avio, K.	nny	+2	tr	Avio (1979)
Fabrikant, R.	nyn	−1	tr	Fabrikant (1979)
Fujii, E.; Mak, J.	nyy	0	tr	Fujii and Mak (1979)
Greenberg, D. et al.	nyy	−2	tr	Greenberg et al. (1979)
Hakim, S. et al.	nyy	−2	tr	Hakim et al. (1979)
Kleck, G.	nny	−1	ah	Kleck (1979)
Otterbein, K.	nny	−1	kr	Otterbein (1979)
Parker, R.; Smith, M.	nny	−1	kr	Parker and Smith (1979)
Peek, C. et al.	ynn	−1	aw	Peek et al. (1979)
Pontell, H.	nyy	+2	aw	Pontell (1979)
Sickles, R. et al.	yyy	−1	tr	Sickles et al. (1979)
Storey, D.	yyn	−1	tr	Storey (1979)

...last page of table B.5 continued

Authors	CT	O	U	Source
Albrecht, H.	yyn	+1	aw	Albrecht (1980)
Bailey, W.	nny	+1	aw	Bailey (1980a)
Bailey, W.	nny	−1	aw	Bailey (1980b)
Brier, S.; Fienberg, S.	yyy	+1	tr	Brier and Fienberg (1980)
Bryjak, G.	nyn	+1	aw	Bryjak (1980)
Burkett, S.; Carrithers, W.	ynn	−1	ah	Burkett and Carrithers (1980)
Fujii, E.; Mak, J.	nyy	0	tr	Fujii and Mak (1980)
Goldberg, I.; Nold, F.	nyn	−2	tr	Goldberg and Nold (1980)
Grasmick, H.; Bryjak, G.	nyy	−1	aw	Grasmick and Bryjak (1980)
Hakim, S.	nyn	0	tr	Hakim (1980)
Hoenack, S.; Weiler, W.	nny	−1	tr	Hoenack and Weiler (1980)
Huff, R.; Stahura, J.	nyy	+2	ah	Huff and Stahura (1980)
Humphries, D.; Wallace, D.	nyy	+2	ah	Humphries and Wallace (1980)
Kirchner, R. et al.	nyn	+1	kr	Kirchner et al. (1980)
Loftin, C.	nny	+1	aw	Loftin (1980)
Phillips, D.	nny	−1	ah	Phillips (1980)
Witte, A.	yyy	−1	tr	Witte (1980)
Wolpin, K.	nyn	−2	tr	Wolpin (1980)
Cloninger, D.	ynn	−2	tr	Cloninger (1981)
Corman, H.	nyn	−1	tr	Corman (1981)
Ehrlich, I.	nyy	0	tr	Ehrlich (1981)
Furlong, W.; Mehay, S.	nyy	−2	tr	Furlong and Mehay (1981)
Gabor, T.	nyn	+1	kr	Gabor (1981)
Jacob, H.; Rich, M.	yyy	+2	ah	Jacob and Rich (1981)
Miranne, A.	nyn	+1	aw	Miranne (1981)
Pierce, G.; Bowers, W.	nyy	+1	ah	Pierce and Bowers (1981)
Scott, W.; Grasmick, H.	nyn	−2	tr	Scott and Grasmick (1981)
Williams, K.; Gibbs, J.	yyy	−1	ah	Williams and Gibbs (1981)
Bailey, W.	nny	+2	ah	Bailey (1982)
Biles, D.	yyy	+1	ah	Biles (1982)
Bishop, D.	ynn	0	aw	Bishop (1983)
Chilton, R.	yyy	−1	aw	Chilton (1982)
Grasmick, H.; Scott, W.	nyn	−1	aw	Grasmick and Scott (1982)
Greenberg, D.; Kessler, R.	nyy	+1	ah	Greenberg and Kessler (1982a)
Greenberg, D.; Kessler, R.	nyy	−1	tr	Greenberg and Kessler (1982b)
Gross, M.; Hakim, S.	yyn	+2	tr	Gross and Hakim (1982)
Hannan, T.	nyn	−1	ah	Hannan (1982)
Jensen, G.; Stitt, B.	yyn	−1	aw	Jensen and Stitt (1982)
Leamer, E.	nny	0	tr	Leamer (1982)
Loftin, C.; McDowall, D.	nyy	+2	tr	Loftin and McDowall (1982)
Medoff, M.; Magaddino, J.	nny	+2	kr	Medoff and Magaddino (1982)
Meier, R.	ynn	+1	ah	Meier (1982)
Minor, W.; Harry, J.	yyn	+1	aw	Minor and Harry (1982)
Myers, S.	nyn	−2	tr	Myers (1982)
Parilla, P.	nyn	−1	aw	Parilla (1982)
Paternoster, R. et al.	yyn	+1	aw	Paternoster et al. (1982a)

…last page of table B.5 continued

Authors	CT	O	U	Source
Paternoster, R. et al.	yyn	−1	aw	Paternoster et al. (1982b)
Rankin, J.; Wells, L.	yyy	−1	aw	Rankin and Wells (1982)
Saltzman, L. et al.	yyn	−1	aw	Saltzman et al. (1982)
Sesnowitz, M.; Hexter, J.	nyn	−2	tr	Sesnowitz and Hexter (1982)
Spicer, M.; Thomas, E.	nyn	−2	tr	Spicer and Thomas (1982)
Stack, S.	nyn	+1	aw	Stack (1982)
Wärneryd, K.; Walerud, B.	nyn	−1	ah	Wärneryd and Walerud (1982)
Bailey, W.	nny	+1	ah	Bailey (1983)
Dölling, D.	yyy	−2	mw	Dölling (1983)
Forst, B.	nny	+1	ah	Forst (1983)
Ghali, M. et al.	yyy	−1	tr	Ghali et al. (1983)
Grasmick, H. et al.	nyy	−1	aw	Grasmick et al. (1983)
Greenberg, D. et al.	nyy	+2	ah	Greenberg et al. (1983)
Hollinger, R.; Clark, J.	nyn	−1	aw	Hollinger and Clark (1983)
Houston, J.	nyn	0	tr	Houston (1983)
Kohfeld, C.	nyy	−1	aw	Kohfeld (1983)
Layson, S.	nny	−1	tr	Layson (1983)
Leamer, E.	nny	0	tr	Leamer (1983)
Liu, Y.; Bee, R.	nyn	−1	tr	Liu and Bee (1983)
Low, S.; McPheters, L.	yyy	0	tr	Low and McPheters (1983)
McFarland, S.	nny	+2	aw	McFarland (1983)
Norström, T.	ynn	+2	ah	Norström (1983)
Paternoster, R. et al.	nyn	+1	mw	Paternoster et al. (1983a)
Paternoster, R. et al.	yyn	+1	aw	Paternoster et al. (1983b)
Votey, H.; Shapiro, P.	ynn	−2	tr	Votey and Shapiro (1983)
Willis, K.	nyy	−2	tr	Willis (1983)
Zedlewski, E.	yyn	−1	aw	Zedlewski (1983)
Bishop, D.	yyy	−1	aw	Bishop (1984a)
Bishop, D.	nyy	−1	aw	Bishop (1984b)
Clark, A.	nyn	−2	aw	Clark (1984)
Decker, S.; Kohfeld, C.	nny	+2	aw	Decker and Kohfeld (1984)
Epple, D.; Visscher, M.	nyn	−2	tr	Epple and Visscher (1984)
Hilton, M.	ynn	−1	aw	Hilton (1984)
Krylo, D.	nyn	+1	aw	Krylo (1985)
McCormick, R.; Tollison, R.	ynn	−2	tr	McCormick and Tollison (1984)
Meier, R. et al.	ynn	−1	kr	Meier et al. (1984)
Pestello, F.	yyy	−1	aw	Pestello (1984)
Stack, S.	nyn	+2	aw	Stack (1984)
Swan, P.	ynn	−1	aw	Swan (1984)
Sykes, G.	ynn	+2	aw	Sykes (1984)
Votey, H.	yyy	−1	tr	Votey (1984a)
Votey, H.	ynn	−2	tr	Votey (1984b)
Withers, G.	yyy	−1	tr	Withers (1984)
Zimring, F.	nny	−1	ah	Zimring (1984)
Benjamini, Y.; Maital, S.	nyn	+1	tr	Benjamini and Maital (1985)
Friedland, N.	nyn	−1	kr	Friedland (1985)

. . . last page of table B.5 continued

Authors	CT	O	U	Source
Isachsen, A. et al.	yyn	−1	tr	Isachsen et al. (1985)
Jackson, Betty R.; Jones, S.	nyn	−1	ah	Jackson and Jones (1985)
Killias, M.	ynn	−1	aw	Killias (1985)
Layson, S.	nny	−2	tr	Layson (1985)
Layton, A.	ynn	−2	tr	Layton (1983)
McManus, W.	nny	0	tr	McManus (1985)
Montmarquette, C. et al.	yyn	−1	aw	Montmarquette et al. (1985)
Sheley, J.; Bailey, K.	ynn	−1	aw	Sheley and Bailey (1985)
Spicer, M.; Hero, R.	nyn	−1	tr	Spicer and Hero (1985)
Williams, F.	ynn	+2	aw	Williams (1985)
Wilson, R.; Jonah, B.	ynn	−1	tr	Wilson and Jonah (1985)
Witte, A.; Woodbury, D.	nyn	+1	ah	Witte and Woodbury (1985)
Berger, D.; Snortum, J.	ynn	+2	tr	Berger and Snortum (1986)
Blau, P.; Golden, R.	yyy	−1	ah	Blau and Golden (1986)
Bursik, R.; Baba, Y.	ynn	+1	ah	Bursik and Baba (1986)
Gyimah-Brempong, K.	nyy	0	tr	Gyimah-Brempong (1986)
Miller, J.; Anderson, A.	nyn	−1	ah	Miller and Anderson (1986)
Paternoster, R.; Iovanni, L.	nyn	+1	mw	Paternoster and Iovanni (1986)
Pogue, T.	nyn	+1	tr	Pogue (1986)
Stafford, M. et al.	yny	−2	tr	Stafford et al. (1986)
Viscusi, W.	nyn	−2	tr	Viscusi (1986)
Ward, D. at al.	ynn	−1	kr	Ward et al. (1986)
Watson, R.	ynn	−1	aw	Watson (1986)
Webley, P.; Halstead, S.	nyn	0	ah	Webley and Halstead (1986)
Becker, W. et al.	nyn	−1	ah	Becker et al. (1987)
Berlitz, C. et al.	yyy	+1	aw	Berlitz et al. (1987)
Cloninger, D.	nny	−2	tr	Cloninger (1987)
Cohen, L.; Land, K.	nyy	−1	ah	Cohen and Land (1987)
Cohen, M.	nyn	−2	tr	Cohen (1987)
Corman, H. et al.	nyn	−2	tr	Corman et al. (1987)
Demers, D.; Lundman, R.	ynn	−1	kr	Demers and Lundman (1987)
Ehrlich, I.; Brower, G.	nny	−2	tr	Ehrlich and Brower (1987)
Good, D.; Pirog-Good, M.	ynn	0	tr	Good and Pirog-Good (1987)
Howsen, R.; Jarrell, S.	nyn	−1	tr	Howsen and Jarrell (1987)
Kalfus, G. et al.	yyn	−1	aw	Kalfus et al. (1987)
Lott, J.	yyy	0	tr	Lott (1987)
McCarthy, P.; Oesterle, W.	ynn	−2	tr	McCarthy and Oesterle (1987)
Miranne, A.; Gray, L.	nyn	+1	kr	Miranne and Gray (1987)
Rabow, J. et al.	ynn	−1	tr	Rabow et al. (1987)
Ross, L.	ynn	−2	tr	Ross (1987a)
Ross, L.	ynn	−1	aw	Ross (1987b)
Schumann, K. et al.	yyy	+1	aw	Schumann et al. (1987)
Smith, D.; Paternoster, R.	yyn	+2	ah	Smith and Paternoster (1987)
Stack, S.	nny	−1	aw	Stack (1987)
Voas, R.; Hause, J.	ynn	−2	tr	Voas and Hause (1987)
Wilkinson, J.	ynn	+2	tr	Wilkinson (1987)

...last page of table B.5 continued

Authors	CT	O	U	Source
Avio, K.	nny	+1	tr	Avio (1988)
Beron, K. et al.	nyn	−1	tr	Beron et al. (1988)
Burnell, J.	nyn	+1	tr	Burnell (1988)
Cover, J.; Thistle, P.	nny	0	tr	Cover and Thistle (1988)
Devine, J. et al.	nyy	−1	ah	Devine et al. (1988)
Dubin, J.; Wilde, L.	nyn	+1	ah	Dubin and Wilde (1988)
Hessing, D. et al.	nyn	0	tr	Hessing et al. (1988)
Hite, P.	nyn	+2	ah	Hite (1988)
Maghsoodloo, S. et al.	ynn	−1	tr	Maghsoodloo et al. (1988)
Merriman, D.	nny	−1	kr	Merriman (1988)
Paternoster, R.	yyn	−1	ah	Paternoster (1988)
Shore, E.; Maguin, E.	ynn	−2	tr	Shore and Maguin (1988)
Smith, D.	ynn	−1	aw	Smith (1988)
Stevans, L.	nyn	−2	tr	Stevans (1988)
Bailey, W.; Peterson, R.	nny	+1	ah	Bailey and Peterson (1989)
Bönitz, D.	yyy	+1	aw	Bönitz (1989)
Cameron, S.	nny	−1	tr	Cameron (1989)
Chressanthis, G.	nny	−2	tr	Chressanthis (1989)
Craig, S.; Heikkila, E.	nyn	−2	tr	Craig and Heikkila (1989)
Friedman, J. et al.	ynn	−1	mw	Friedman et al. (1989)
Gillis, A.	nyy	0	ah	Gillis (1989)
Goel, R.; Rich, D.	ynn	−2	tr	Goel and Rich (1989)
Green, D.	ynn	+1	aw	Green (1989a)
Green, D.	ynn	−2	aw	Green (1989b)
Haque, M.; Cameron, M.	ynn	+2	aw	Haque and Cameron (1989)
von Hofer, H.; Tham, H.	nyn	+2	aw	von Hofer and Tham (1989)
Keane, C. et al.	ynn	+1	aw	Keane et al. (1989)
Klepper, S.; Nagin, D.	nyn	0	ah	Klepper and Nagin (1989a)
Klepper, S.; Nagin, D.	nyn	0	kr	Klepper and Nagin (1989b)
Klepper, S.; Nagin, D.	nyn	+2	aw	Klepper and Nagin (1989c)
McAleer, M.; Veall, M.	nny	+2	tr	McAleer and Veall (1989)
Muller, A.	ynn	−1	aw	Muller (1989)
Paternoster, R.	yyn	0	aw	Paternoster (1989a)
Paternoster, R.	yyn	+2	ah	Paternoster (1989b)
Saffer, H.; Chaloupka, F.	ynn	−2	tr	Saffer and Chaloupka (1989)
Snortum, J.; Berger, D.	ynn	−1	tr	Snortum and Berger (1989)
Stalans, L. et al.	nyn	−1	aw	Stalans et al. (1989)
Trumbull, W.	yyy	−2	tr	Trumbull (1989)
Zador, P. et al.	ynn	−1	ah	Zador et al. (1989)
Alm, J. et al.	nyn	0	tr	Alm et al. (1990a)
Alm, J. et al.	nyn	+2	tr	Alm et al. (1990b)
Bailey, W.	nny	+1	aw	Bailey (1990)
Bursik, R. et al.	nyn	+2	aw	Bursik et al. (1990)
Caudill, B. et al.	ynn	+1	tr	Caudill et al. (1990)
Chressanthis, G.; Grimes, P.	nyy	−2	tr	Chressanthis and Grimes (1990)
Clark, D.; Cosgrove, J.	nny	+1	tr	Clark and Cosgrove (1990)

. . . last page of table B.5 continued

Authors	CT	O	U	Source
Cloninger, D.	yyn	−2	tr	Cloninger (1990)
Corman, H.; Joyce, T.	nyy	0	aw	Corman and Joyce (1990)
Decker, S.; Kohfeld, C.	nyn	−1	tr	Decker and Kohfeld (1990a)
Decker, S.; Kohfeld, C.	nny	+1	ah	Decker and Kohfeld (1990b)
Field, S.	yyy	0	tr	Field (1990)
Gartner, R.	nny	−1	ah	Gartner (1990)
Gibbs, J.; Firebaugh, G.	yyy	−1	aw	Gibbs and Firebaugh (1990)
Grasmick, H.; Bursik, R.	yyn	−1	aw	Grasmick and Bursik (1990)
Grogger, J.	nny	+2	tr	Grogger (1990)
Jarrell, S.; Howsen, R.	nyy	−1	aw	Jarrell and Howsen (1990)
Karstedt-Henke, S.	nyy	−1	aw	Karstedt-Henke (1991)
Legge, J.	ynn	−1	aw	Legge (1990)
Magat, W.; Viscusi, W.	yyn	−2	tr	Magat and Viscusi (1990)
Mikesell, J.; Pirog-Good, M.	nyn	+1	tr	Mikesell and Pirog-Good (1990)
Ross, L.; Voas, R.	ynn	−1	kr	Ross and Voas (1990)
Ross, L. et al.	ynn	+2	tr	Ross et al. (1990)
Schumann, K.; Kaulitzki, R.	yyy	+1	aw	Schumann and Kaulitzki (1991)
Soper, J.; Thompson, L.	ynn	+2	tr	Soper and Thompson (1990)
Stack, S.	nny	−1	ah	Stack (1990)
Virén, M.	nyn	−2	tr	Virén (1990)
Bailey, W.	nyy	+2	aw	Bailey (1991)
Braithwaite, J.; Makkai, T.	yyn	+1	tr	Braithwaite and Makkai (1991)
Cappell, C.; Sykes, G.	nyy	+1	ah	Cappell and Sykes (1991)
Collins, J.; Plumlee, R.	nyn	+2	aw	Collins and Plumlee (1991)
Evans, W. et al.	ynn	+1	tr	Evans et al. (1991)
Furlong, W.	nyn	−2	tr	Furlong (1991)
Laycock, G.	nyn	−2	tr	Laycock (1991)
Mann, R. et al.	ynn	0	tr	Mann et al. (1991)
Nagin, D.; Paternoster, R.	yyn	−1	ah	Nagin and Paternoster (1991)
Peterson, R.; Bailey, W.	nny	+1	ah	Peterson and Bailey (1991)
Thurman, Q.	nyn	+2	tr	Thurman (1991)
Alm, J. et al.	nyn	0	tr	Alm et al. (1992a)
Alm, J. et al.	nyn	−2	tr	Alm et al. (1992b)
Benson, B. et al.	nyn	−2	tr	Benson et al. (1992)
Bunn, D. et al.	yyn	+2	tr	Bunn et al. (1992)
Chamlin, M. et al.	nyn	+1	ah	Chamlin et al. (1992)
Cloninger, D.	yyy	−2	tr	Cloninger (1992)
Erard, B.	nyn	+2	tr	Erard (1992)
Hessing, D. et al.	nyn	+2	ah	Hessing et al. (1992)
McDowall, D. et al.	nyy	0	ah	McDowall et al. (1992)
Meier, K.	nyy	+2	ah	Meier (1992)
Pate, A.; Hamilton, E.	nny	+2	aw	Pate and Hamilton (1992)
Sheffrin, S.; Triest, R.	nyn	−2	tr	Sheffrin and Triest (1992)
Smith, K.	nyn	−2	tr	Smith (1992)
Steenbergen, M. et al.	nyn	+2	tr	Steenbergen et al. (1992)
van Tulder, F.	yyy	−1	ah	van Tulder (1992)

. . . last page of table B.5 continued

Authors	CT	O	U	Source
Ward, S. et al.	nny	−1	aw	Ward et al. (1992)
Williams, K.	nny	−2	mw	Williams (1992)
Burkett, S.; Ward, D.	ynn	−1	aw	Burkett and Ward (1993)
Chaloupka, F. et al.	ynn	−1	tr	Chaloupka et al. (1993)
Cheatwood, D.	nny	+2	aw	Cheatwood (1993)
Ensor, T.; Godfrey, C.	yyy	0	tr	Ensor and Godfrey (1993)
Kenkel, D.	ynn	−2	tr	Kenkel (1993)
Koskela, E.; Virén, M.	nyn	−2	tr	Koskela and Virén (1993)
Neustrom, M.; Norton, W.	ynn	−2	tr	Neustrom and Norton (1993)
Rasmussen, D. et al.	nny	−2	tr	Rasmussen et al. (1993)
Rhee, L.; Zhang, J.	ynn	−1	aw	Rhee and Zhang (1993)
Yu, J.; Liska, A.	nyy	−1	kr	Yu and Liska (1993)
Bailey, W.; Peterson, R.	nny	+2	ah	Bailey and Peterson (1994)
Cloninger, D.	nyy	−1	tr	Cloninger (1994)
Cochran, J. et al.	nny	+2	ah	Cochran et al. (1994)
Cornwell, C.; Trumbull, W.	yyy	0	tr	Cornwell and Trumbull (1994)
Denq, F. et al.	yyy	+1	aw	Denq et al. (1994)
Eide, E. et al.	yyy	−2	tr	Eide et al. (1994)
Homel, R.	ynn	−1	ah	Homel (1994)
Jensen, E.; Metsger, L.	nyy	−1	kr	Jensen and Metsger (1994)
Koskela, E.; Virén, M.	nyn	−2	tr	Koskela and Virén (1994)
Legge, J.; Park, J.	ynn	+2	kr	Legge and Park (1994)
Marvell, T.; Moody, C.	yyy	+2	tr	Marvell and Moody (1994)
Mullahy, J.; Sindelar, J.	ynn	−2	tr	Mullahy and Sindelar (1994)
Paternoster, R.; Nagin, D.	yyy	−1	aw	Paternoster and Nagin (1994)
Niskanen, W.	nyy	+2	aw	Niskanen (1994)
Petee, T. et al.	yyn	−1	kr	Petee et al. (1994)
Pyle, D.; Deadman, D.	nyn	−2	tr	Pyle and Deadman (1994)
Rogers, P.; Schoenig, S.	ynn	−1	tr	Rogers and Schoenig (1994)
Sloan, F.; Githens, P.	ynn	−2	tr	Sloan and Githens (1994)
Sloan, F. et al.	yny	−1	tr	Sloan et al. (1994)
Sollars, D. et al.	yyn	+1	aw	Sollars et al. (1994)
Tauchen, H. et al.	yyy	−2	tr	Tauchen et al. (1994)
Virén, M.	yyn	−2	tr	Virén (1994)
Wieczorek, W. et al.	ynn	+1	aw	Wieczorek et al. (1994)
Anderson, E.; Talley, W.	nyn	+1	tr	Anderson and Talley (1995)
Andreoni, J.	yyy	+2	tr	Andreoni (1995)
Brumm, H.; Cloninger, D.	nny	−2	tr	Brumm and Cloninger (1995a)
Brumm, H.; Cloninger, D.	nyy	−2	tr	Brumm and Cloninger (1995b)
Elis, L.; Simpson, S.	ynn	+1	aw	Elis and Simpson (1995)
Gertz, M.; Gould, L.	yyy	−1	ah	Gertz and Gould (1995)
Hull, B.; Bold, F.	yyy	+2	mw	Hull and Bold (1995)
Johnson, D.; Fell, J.	ynn	−2	tr	Johnson and Fell (1995)
Masih, R.	nyn	+1	tr	Masih (1995)
McDowall, D. et al.	nny	+2	ah	McDowall et al. (1995)
Meera, A.; Jayakumar, M.	yyy	−1	tr	Meera and Jayakumar (1995)

... last page of table B.5 continued

Authors	CT	O	U	Source
Parker, R.	nny	+2	aw	Parker (1995)
Ross, L.; Klette, H.	ynn	+2	tr	Ross and Klette (1995)
Sherman, L.; Rogan, D.	yyy	−1	aw	Sherman and Rogan (1995)
Sloan, F. et al.	ynn	+2	ah	Sloan et al. (1995)
Wong, Y.	yyy	0	tr	Wong (1995)
Allen, R.	nyn	+1	tr	Allen (1996)
Anderson, E.; Diaz, J.	yyy	−1	tr	Anderson and Diaz (1996)
Brumm, H.; Cloninger, D.	nny	−1	ah	Brumm and Cloninger (1996)
Deshapriya, E.; Iwase, N.	ynn	−1	tr	Deshapriya and Iwase (1996)
Entorf, H.	yyy	+1	tr	Entorf (1996)
Hingson, R. et al.	ynn	−1	aw	Hingson et al. (1996)
Hsing, B.	ynn	−1	aw	Hsing (1996)
Levitt, S.	nyy	−1	tr	Levitt (1996)
Mann, R. et al.	ynn	+2	tr	Mann et al. (1996)
Marvell, T.; Moody, C.	nyy	−1	aw	Marvell and Moody (1996)
Mixon, F.; Mixon, D.	nyn	+1	tr	Mixon and Mixon (1996)
Paternoster, R.; Simpson, S.	nyn	+1	ah	Paternoster and Simpson (1996)
Pommerehne, W.; Weck-Hannemann, H.	nyn	+1	ah	Pommerehne and Weck-Hannemann (1996)
Reilly, B.; Witt, R.	nyn	−1	tr	Reilly and Witt (1996)
Ruhm, C.	ynn	+2	tr	Ruhm (1996)
Berger, D.; Marelich, W.	ynn	−2	tr	Berger and Marelich (1997)
Bodman, P.; Maultby, C.	nyn	−2	tr	Bodman and Maultby (1997)
Bosco, L.; Mittone, L.	nyn	+2	tr	Bosco and Mittone (1997)
Foglia, W.	yyy	−1	aw	Foglia (1997)
Goodman, D.	nyy	−2	tr	Goodman (1997)
Grau, M.; Groves, T.	nyn	0	tr	Grau and Groves (1997)
Johnson, J.; Lesniak-Karpiak, K.	nyn	−1	aw	Johnson and Lesniak-Karpiak (1997)
Kaplan, H.; Damphousse, K.	nyn	+1	aw	Kaplan and Damphousse (1997)
Koskela, E.; Virén, M.	nyn	−2	tr	Koskela and Virén (1997)
Levitt, S.	nyy	−1	tr	Levitt (1997b)
Lott, J.; Mustard, D.	nyy	−1	tr	Lott and Mustard (1997)
Marselli, R.; Vannini, M.	nyy	−1	tr	Marselli and Vannini (1997)
Marvell, T.; Moody, C.	nyy	−1	aw	Marvell and Moody (1997)
McGeorge, J.; Aitken, C.	ynn	+2	tr	McGeorge and Aitken (1997)
Merrifield, J.	nyn	−2	tr	Merrifield (1997)
Mui, H.; Ali, M.	ynn	0	tr	Mui and Ali (1997)
Olson, D.	nyy	−1	aw	Olson (1997)
Raftery, A. et al.	yyy	0	tr	Raftery et al. (1997)
Sirakaya, E.; Uysal, M.	ynn	+1	tr	Sirakaya and Uysal (1997)
Thomson, E.	nyn	+2	aw	Thomson (1997)
Voas, R. et al.	ynn	0	tr	Voas et al. (1997)
Zhang, J.	nyn	+1	tr	Zhang (1997)
Ayres, I.; Levitt, S.	nyy	−2	tr	Ayres and Levitt (1998)
Bailey, W.	nny	+1	ah	Bailey (1998)
Balbo, M.; Posadas, J.	yyy	0	tr	Balbo and Posadas (1998)
Baron, S.; Kennedy, L.	nyy	0	aw	Baron and Kennedy (1998)

…last page of table B.5 continued

Authors	CT	O	U	Source
Benson, B. et al.	yyy	−2	tr	Benson et al. (1998)
Black, D.; Nagin, D.	nyy	+1	aw	Black and Nagin (1998)
Borack, J.	ynn	−2	tr	Borack (1998)
Bronars, S.; Lott, J.	nyy	−1	ah	Bronars and Lott (1998)
Chambouleyron, A.; Willington, M.	ynn	−2	tr	Chambouleyron and Willington (1998)
Dezhbakhsh, H.; Rubin, P.	nyy	+1	tr	Dezhbakhsh and Rubin (1998)
DiPasquale, D.; Glaeser, E.	ynn	−1	tr	DiPasquale and Glaeser (1998)
Entorf, H.; Spengler, H.	yyy	−2	tr	Entorf and Spengler (2000)
Fajnzylber, P. et al.	nyy	−2	tr	Fajnzylber et al. (1998)
Fishman, G. et al.	nyn	−2	aw	Fishman et al. (1998)
Foss, R. et al.	ynn	+2	tr	Foss et al. (1998)
Goel, R.; Nelson, M.	ynn	0	tr	Goel and Nelson (1998)
Hale, C.	nyn	−1	ah	Hale (1998)
Kuperan, K.; Sutinen, J.	ynn	+1	ah	Kuperan and Sutinen (1998)
Levitt, S.	nyy	−2	tr	Levitt (1998a)
Levitt, S.	nyy	−1	tr	Levitt (1998c)
Lott, J.	nyy	−2	tr	Lott (1998)
Ludwig, J.	nny	+2	tr	Ludwig (1998)
Piquero, A.; Paternoster, R.	ynn	+1	ah	Piquero and Paternoster (1998)
Sigman, H.	nyn	−1	tr	Sigman (1998)
Taxman, F.; Piquero, A.	ynn	+2	tr	Taxman and Piquero (1998)
Vingilis, E. et al.	ynn	+1	tr	Vingilis et al. (1988)
Witt, R. et al.	nyn	+1	tr	Witt et al. (1998)
Becsi, Z.	yyy	−2	tr	Becsi (1999)
Benson, B. et al.	ynn	−1	aw	Benson et al. (1999)
Cherry, T.	nyy	0	tr	Cherry (1999)
Cochran, J. et al.	yyn	+1	ah	Cochran et al. (1999)
Curti, H.	nyy	−1	tr	Curti (1999)
Doyle, J. et al.	nyy	−1	tr	Doyle et al. (1999)
Ehrlich, I.; Zhiqiang, L.	nyy	−2	tr	Ehrlich and Zhiqiang (1999)
Farrelly, M. et al.	ynn	0	tr	Farrelly et al. (1999)
Gawande, K.; Wheeler, T.	yyn	−2	tr	Gawande and Wheeler (1999)
Gius, M.	nyy	+2	tr	Gius (1999)
Grosvenor, D. et al.	ynn	+1	tr	Grosvenor et al. (1999)
Hale, C.	nyn	+1	tr	Hale (1999)
Hatcher, A. et al.	ynn	−2	tr	Hatcher et al. (1999)
Curti, H.	yyn	−2	tr	Curti (1984)
Kessler, D.; Levitt, S.	nyy	−1	tr	Kessler and Levitt (1998)
Lynch, M.	yyy	+2	aw	Lynch (1999)
MacDonald, J.	yny	0	tr	MacDonald (1999)
Mehay, S.; Pacula, R.	ynn	0	tr	Mehay and Pacula (1999)
Mocan, H.; Rees, D.	yyy	−1	tr	Mocan and Rees (2005)
Olson, M.	nyn	−1	aw	Olson (1999)
Papps, K.; Winkelmann, R.	yyy	+2	tr	Papps and Winkelmann (1999)
Parker, J.; Atkins, R.	nyn	+2	tr	Parker and Atkins (1999)
Ralston, R.	nyn	0	tr	Ralston (1999)

... last page of table B.5 continued

Authors	CT	O	U	Source
Retting, R. et al.	ynn	−2	tr	Retting et al. (1999)
Sorensen, J. et al.	nny	+2	aw	Sorensen et al. (1999)
Spencer, D.	yyy	0	aw	Spencer (1999)
Thomson, E.	nny	+2	aw	Thomson (1999)
Tibbetts, S.	nyn	+1	aw	Tibbetts (1999)
van Tulder, F.; van der Torre, A.	yyy	−1	tr	van Tulder and van der Torre (1999)
Vinod, H.	ynn	−2	tr	Vinod (1999)
Witt, R. et al.	yyn	+1	aw	Witt et al. (1999)
Benson, B. et al.	ynn	−1	tr	Benson et al. (2000)
Cerro, A.; Meloni, O.	yyy	−2	tr	Cerro and Meloni (2000)
Cochran, J.; Chamlin, M.	nny	+2	ah	Cochran and Chamlin (2000)
Corman, H.; Mocan, H.	nyy	−2	tr	Corman and Mocan (2000)
Deadman, D.; Pyle, D.	nyn	−2	tr	Deadman and Pyle (2000)
Diez-Ticio, A.	nyn	−2	tr	Diez-Ticio (2000)
Fajnzylber, P. et al.	nyy	0	tr	Fajnzylber et al. (2000)
Fehr, E.; Gächter, S.	ynn	−2	tr	Fehr and Gächter (2000)
Charmichael, F.; Ward, R.	yyy	+1	tr	Carmichael and Ward (2000)
Giacopassi, D.; Forde, D.	nny	+1	aw	Giacopassi and Forde (2000)
Gneezy, U.; Rustichini, A.	ynn	+2	tr	Gneezy and Rustichini (2000)
Kelly, M.	nyy	0	tr	Kelly (2000)
Mast, B. et al.	ynn	0	tr	Mast et al. (2000)
Pudney, S. et al.	nyn	−2	tr	Pudney et al. (2000a)
Pudney, S. et al.	nyn	−1	tr	Pudney et al. (2000b)
Resignato, A.	nyy	−2	tr	Resignato (2000)
Spelman, W.	yyy	−1	tr	Spelman (2000)
Stolzenberg, L.; D'Alessio, S.	yyy	0	mw	Stolzenberg and D'Alessio (2000)
Stout, E. et al.	ynn	0	tr	Stout et al. (2000)
Weber, J.; Crew, R.	nyn	−1	tr	Weber and Crew (2000)
Whetten-Goldstein, K. et al	ynn	+1	tr	Whetten-Goldstein et al. (2000)
Witt, R.; Witte, A.	yyy	−1	tr	Witt and Witte (2000)
Yu, J.	ynn	+1	tr	Yu (2000)
Bar-Ilan, A.; Sacerdote, B.	ynn	−2	tr	Bar-Ilan and Sacerdote (2001)
Benson, B. et al.	nyy	+1	tr	Benson et al. (2001)
Benson, B.; Mast, B.	nyy	+1	aw	Benson and Mast (2001)
Braga, A. et al.	nny	−1	aw	Braga et al. (2001)
Cherry, T.	yyy	−2	tr	Cherry (2001)
Cloninger, D.; Marchesini, R.	nyy	−2	tr	Cloninger and Marchesini (2001)
Cummings, R. et al.	nyn	+1	aw	Cummings et al. (2001)
DeSimone, J.	nyy	−2	tr	DeSimone (2001)
Dezhbakhsh, H. et al.	nny	−2	tr	Dezhbakhsh et al. (2001)
Fernández, C. et al.	yyy	0	tr	Fernández et al. (2001a)
Gunnison, E.	nyy	+2	aw	Gunnison (2002)
Shachmurove, Y. et al.	nyn	−2	tr	Shachmurove et al. (2001)
Kelling, G.; Sousa, W.	nyy	0	tr	Kelling and Sousa (2001)
Kenkel, D.; Koch, S.	ynn	+2	tr	Kenkel and Koch (2001)
Levitt, S.	nyy	0	tr	Levitt (2001)

...last page of table B.5 continued

Authors	CT	O	U	Source
Levitt, S.; Lochner, L.	yyy	−1	tr	Levitt and Lochner (2001)
Levitt, S.; Porter, J.	ynn	0	tr	Levitt and Porter (2001)
Liang, F. et al.	yyy	0	tr	Liang et al. (2001)
Lochner, L.	yyn	−2	tr	Lochner (2001)
Lott, J.; Whitley, J.	nyy	−2	ah	Lott and Whitley (2001)
Luiz, J.	yyy	+2	ah	Luiz (2001)
Marvell, T.; Moody, C.	nyy	+2	aw	Marvell and Moody (2001)
McGarrell, E. et al.	nyy	+1	ah	McGarrell et al. (2001)
Miron, J.	nny	+2	ah	Miron (2001)
Nagin, D.; Pogarsky, G.	ynn	−1	aw	Nagin and Pogarsky (2001)
Olson, D.; Maltz, M.	nny	−1	ah	Olson and Maltz (2001)
Parsley, J.	nyy	−2	tr	Parsley (2001)
Plassmann, F.; Tideman, N.	nyy	+1	ah	Plassmann and Tideman (2001)
Winter-Ebmer, R.; Raphael, S.	nyy	−1	ah	Winter-Ebmer and Raphael (2001)
Scribner, R.; Cohen, D.	ynn	−1	aw	Scribner and Cohen (2001)
Slemrod, J. et al.	nyn	+2	aw	Slemrod et al. (2001)
Sullivan, K. et al.	ynn	+2	tr	Sullivan et al. (2001)
Virén, M.	nyn	−2	tr	Virén (2001)
Yunker, J.	nny	+2	aw	Yunker (2001)
Allen, W.	ynn	0	tr	Allen (2002)
Andrienko, Y.	nyy	−2	tr	Andrienko (2002)
Cherry, T.; List, J.	nyy	0	tr	Cherry and List (2002)
Cook, P.; Ludwig, J.	nyn	0	tr	Cook and Ludwig (2002)
Corman, H.; Mocan, H.	nyy	−2	tr	Corman and Mocan (2002)
DeFina, R.; Arvanites, T.	nyy	0	aw	DeFina and Arvanites (2002)
Entorf, H.; Spengler, H.	nyn	−2	tr	Entorf and Spengler (2002)
Entorf, H.; Winker, P.	yyy	+1	tr	Entorf and Winker (2002)
Fajnzylber, P. et al.	nyy	0	tr	Fajnzylber et al. (2002a)
Fajnzylber, P. et al.	nyy	−1	tr	Fajnzylber et al. (2002b)
Frey, B.; Feld, L.	nyn	0	tr	Frey and Feld (2002)
Weinberg, B. et al.	yyy	0	tr	Weinberg et al. (2002)
Grasmick, H.; Kobayashi, E.	ynn	+1	ah	Grasmick and Kobayashi (2002)
Hansen, K.; Machin, S.	yyy	+1	tr	Hansen and Machin (2002)
Kaminski, R.; Marvell, T.	nny	+2	aw	Kaminski and Marvell (2002)
Kovandzic, T.; Sloan, J.	nyy	0	aw	Kovandzic and Sloan (2002)
Leung, A.	nyy	+1	tr	Leung (2002)
Levitt, S.	ynn	0	tr	Levitt (2002a)
Levitt, S.	nyy	−2	tr	Levitt (2002b)
Mann, R. et al.	ynn	−1	tr	Mann et al. (2002)
McCrary, J.	nyy	−1	tr	McCrary (2002)
Pfeiffer, M.; Gelau, C.	ynn	−1	aw	Pfeiffer and Gelau (2002)
Piquero, A.; Pogarsky, G.	ynn	−1	ah	Piquero and Pogarsky (2002)
Pogarsky, G.	ynn	−1	aw	Pogarsky (2002)
Shepherd, J.	nyy	−1	aw	Shepherd (2002a)
Shepherd, J.	nyy	+1	ah	Shepherd (2002b)
Stafford, S.	ynn	−1	tr	Stafford (2002)

...last page of table B.5 continued

Authors	CT	O	U	Source
West, A.	nny	+2	ah	West (2002)
Williams, J.; Sickles, R.	nyn	−1	tr	Williams and Sickles (2002)
Anderson, L.; Stafford, S.	ynn	−2	tr	Anderson and Stafford (2003)
Brezina, T.; Piquero, A.	ynn	+1	ah	Brezina and Piquero (2003)
Buonanno, P.	yyy	−1	tr	Buonanno (2003)
Deadman, D.	nyn	−2	tr	Deadman (2003)
DeSimone, J.; Farrelly, M.	ynn	−2	tr	DeSimone and Farrelly (2003)
Dezhbakhsh, H.; Shepherd, J.	nyy	0	tr	Dezhbakhsh and Shepherd (2003)
Elffers, H. et al.	nyn	−1	aw	Elffers et al. (2003)
Entorf, H.	nyy	0	tr	Entorf (2003)
Funk, P.; Kugler, P.	nyn	−2	tr	Funk and Kugler (2003a)
Funk, P.; Kugler, P.	nyn	−1	tr	Funk and Kugler (2003b)
Gainey, R.; Payne, B.	ynn	0	aw	Gainey and Payne (2003)
Heckelman, J.; Yates, A.	ynn	+2	tr	Heckelman and Yates (2003)
Ihlanfeldt, K.	nyy	0	tr	Ihlanfeldt (2003)
Katz, L. et al.	nyy	0	ah	Katz et al. (2003)
Kovandzic, T.; Marvell	nyy	+2	ah	Kovandzic and Marvell (2003)
Lochner, L.	yyy	−1	tr	Lochner (2003)
Mann, R. et al.	ynn	−1	tr	Mann et al. (2003)
Mauser, G.; Maki, D.	nyn	−1	tr	Mauser and Maki (2003)
Mocan, H.; Gittings, K.	nny	+1	aw	Mocan and Gittings (2003)
Mustard, D.	nyy	−1	tr	Mustard (2003)
Nagin, D.; Pogarsky, G.	nyn	−1	aw	Nagin and Pogarsky (2003)
Neumayer, E.	nny	+2	ah	Neumayer (2003)
Park, C.; Hyun, J.	nyn	−2	tr	Park and Hyun (2003)
Richardson, L.; Houston, D.	ynn	+2	tr	Richardson (2003)
Ritsema, C. et al.	nyn	+1	tr	Ritsema et al. (2003)
Sridharan, S. et al.	nyy	+1	tr	Sridharan et al. (2003)
Tao, H.	nyn	−1	tr	Tao (2004)
Tittle, C. et al.	yyy	+1	kr	Tittle et al. (2003)
Baker, T. et al.	ynn	−1	ah	Baker et al. (2004)
Bar-Ilan, A.; Sacerdote, B.	ynn	−1	aw	Bar-Ilan and Sacerdote (2004)
Carmichael, S.; Piquero, A.	nny	0	aw	Carmichael and Piquero (2004)
Denny, K. et al.	nyn	−2	tr	Denny et al. (2004)
Dittrich, M.; Markwardt, G.	yyy	−2	tr	Dittrich and Markwardt (2004)
Earnhart, D.	nyn	−1	aw	Earnhart (2004a)
Earnhart, D.	nyn	−2	tr	Earnhart (2004b)
Feld, L.; Frey, B.	nyn	−1	tr	Feld and Frey (2004)
Fischer, J.	yyy	0	tr	Fischer (2004)
Fortin, B. et al.	nyn	−2	tr	Fortin et al. (2004)
French, M. et al.	ynn	−1	aw	French et al. (2004)
Grxhani, K.; Schram, A.	nyn	0	tr	Gërxhani and Schram (2006)
Kovandzic, T. et al.	nyy	+2	aw	Kovandzic et al. (2004)
Zhiqiang, L.	nny	−2	tr	Zhiqiang (2004)
Lochner, L.	nyn	−2	tr	Lochner (2004)
Machin, S.; Meghir, C.	nyn	−2	tr	Machin and Meghir (2004)

…last page of table B.5 continued

Authors	CT	O	U	Source
Mendes, S.	nyn	−2	aw	Mendes (2004)
Nilsson, A.	nyy	0	tr	Nilsson (2004)
Nott, D.; Green, P.	yyy	0	tr	Nott and Green (2004)
Pogarsky, G.	ynn	0	aw	Pogarsky (2004)
Pogarsky, G.; Piquero, A.	ynn	+2	ah	Pogarsky and Piquero (2004)
Shepherd, J.	nny	−2	tr	Shepherd (2004)
Soares, R.	nyy	0	tr	Soares (2004)
Spengler, H.	yyy	−1	tr	Spengler (2004)
Stolzenberg, L.; D'Alessio, S.	nny	+1	aw	Stolzenberg and D'Alessio (2004)
Wenzel, M.	nyn	−1	aw	Wenzel (2004)
Worrall, J.; Pratt, T.	nyy	0	aw	Worrall and Pratt (2004)
Wright, B. et al.	nyy	−1	ah	Wright et al. (2004)
Zimmerman, P.	nny	−1	tr	Zimmerman (2004)
Carmichael, S. et al.	yyn	−1	ah	Carmichael et al. (2005)
Donohue, J.; Wolfers, J.	nny	0	tr	Donohue and Wolfers (2005)
Dugan, L. et al.	ynn	−1	aw	Dugan et al. (2005)
Entorf, H.; Spengler, H.	nyy	+1	tr	Entorf and Spengler (2005)
Gawande, K.; Bohara, A.	nyn	+2	aw	Gawande and Bohara (2005)
Klick, J.; Tabarrok, A.	nyy	−1	ah	Klick and Tabarrok (2005)
May, P.	ynn	+1	ah	May (2005)
McCarthy, B.; Hagan, J.	yyy	+1	ah	McCarthy and Hagan (2005)
Mocan, H.; Bali, T.	nyy	−2	tr	Mocan and Bali (2005)
Moffett, M. et al.	nyn	−2	ah	Moffett et al. (2005)
Papachristos, A. et al.	nny	+1	ah	Papachristos et al. (2005)
Shepard, E.; Blackley, P.	nyy	−1	ah	Shepard and Blackley (2005)
Shimshack, J.; Ward, M.	nyn	+1	aw	Shimshack and Ward (2005)
Tay, R.	ynn	−1	aw	Tay (2005a)
Tay, R.	ynn	−1	aw	Tay (2005b)
Thornton, D. et al.	ynn	+2	aw	Thornton et al. (2005)
Tittle, C.; Botchkovar, E.	yyy	+2	aw	Tittle and Botchkovar (2005)
Wagenaar, A. et al.	ynn	−1	ah	Wagenaar et al. (2005)
Welch, M. et al.	nyn	−1	ah	Welch et al. (2005)
Wilson, D.	ynn	−1	ah	Wilson (2005)
Wilson, J.; Sheffrin, S.	nyn	+1	aw	Wilson and Sheffrin (2005)
Witt, R.	ynn	−2	tr	Witt (2005)
Antia, K. et al.	ynn	−1	aw	Antia et al. (2006)
Friedman, S. et al.	ynn	+2	ah	Friedman et al. (2006)
Harcourt, B.; Ludwig, J.	nyy	+1	ah	Harcourt and Ludwig (2006)
Kim, K. et al.	ynn	+2	aw	Kim et al. (2006)
Matsueda, R. et al.	nyy	−1	ah	Matsueda et al. (2006)

The *Authors*-column contains the first and second name of the authors. *CT* reports the studied type of crime (violent- property- or other; y=yes, n=no). The third column (*O*) ranges from -2 (strong agreement) to +2 (strong disagreement of the deterrence hypothesis); see subsection 3.3.3. The *U*-column refers to the user who entered the study into the data base while the last column displays the reference and year of publication.

Bibliography

Akers, R. L., M. D. Krohn, L.-K. Lonn, and M. Radosevich (1979). "Social Learning And Deviant Behavior: A Specific Test Of A General Theory." *American Sociological Review.* 44 (4), 636–655.

Albrecht, H.-J. (1980). "Die Generalpräventive Effizienz von Strafrechtlichen Sanktionen." in G. Kaiser (ed.), *Empirische Kriminologie: Ein Jahrzehnt Kriminologischer Forschung am Max-Planck-Institut Freiburg i. Br.*, Freiburg im Breisgau, pp. 305–327.

Alcorn, D. S. (1977). "A Social Psychological Perspective of Deterrence: Development And Test Of A Causal Model." PhD dissertation, Brigham Young University, Provo.

Alexander, C. R. and M. A. Cohen (1999). "Why Do Corporations Become Criminals? Ownership, Hidden Actions, And Crime As An Agency Cost." *Journal Of Corporate Finance.* 5 (1), 1–34.

Allen, D. W. (2002). "Crime, Punishment, And Recidivism: Lessons From The National Hockey League." *Journal Of Sports Economics.* 3 (1), 39–60.

Allen, R. C. (1996). "Socioeconomic Conditions And Property Crime: A Comprehensive Review And Test Of The Professional Literature." *American Journal Of Economics And Sociology.* 55 (3), 293–308.

Allison, J. P. (1972). "Economic Factors And The Rate Of Crime." *Land Economics.* 48 (2), 193–196.

Alm, J., B. R. Jackson, and M. McKee (1992a). "Estimating The Determinants Of Taxpayer Compliance With Experimental Data." *National Tax Journal.* 45 (1), 107–114.

——— , G. H. McClelland, and W. D. Schulze (1992b). "Why Do People Pay Taxes?" *Journal Of Public Economics.* 48 (1), 21–38.

——— , M. McKee, and W. Beck (1990a). "Amazing Grace: Tax Amnesties And Compliance." *National Tax Journal.* 43 (1), 23–37.

——— , R. Bahl, and M. N. Murray (1990b). "Tax Structure And Tax Compliance." *Review Of Economics And Statistics.* 72 (4), 603–613.

Anderson, E. A. and J. Diaz (1996). "Using Process Control Chart Techniques To Analyse Crime Rates In Houston, Texas." *Journal Of The Operational Research Society.* 47 (7), 871–881.

Anderson, E. E. and W. K. Talley (1995). "The Oil Spill Size Of Tanker And Barge Accidents: Determinants And Policy Implications." *Land Economics.* 71 (2), 216–228.

Anderson, L. R. and S. L. Stafford (2003). "Punishment In A Regulatory Setting: Experimental Evidence From The VCM." *Journal Of Regulatory Economics.* 24 (1), 91–110.

Anderson, L. S. (1978). "A Longitudinal Study Of The Deterrence Model." PhD dissertation, Florida State University, Tallahassee.

——— , T. G. Chiricos, and G. P. Waldo (1977). "Formal And Informal Sanctions: A Comparison Of Deterrent Effects." *Social Problems.* 25 (1), 103–114.

Andreoni, J. (1995). "Criminal Deterrence In The Reduced Form: A New Perspective On Ehrlich's Seminal Study." *Economic Inquiry.* 33 (3), 476–483.

Andres, P. and M. Spiwoks (2000). "Prognosegütemaße - State Of The Art der statistischen Ex-Post-Beurteilung von Prognosen." Sofia-Studien zur Institutionenanalyse No. 00-1.

Andrienko, Y. (2002). "What Determines Crime In Russian Regions?" Research Project 99–2521, Economics Education And Research Consortium Russia.

Anglin, W. S. (1992). "Mathematics And History." *Mathematical Intelligencer.* 14 (4), 6–12.

Antia, K. D., M. E. Bergen, S. Dutta, and R. J. Fischer (2006). "How Does Enforcement Deter Gray Market Incidence?" *Journal Of Marketing.* 70 (1), 92–106.

Antony, J. and H. Entorf (2003). "Zur Gültigkeit der Abschreckung im Sinne der ökonomischen Theorie der Kriminalität: Grundzüge einer Metastudie." in H. J. Albrecht and H. Entorf (eds.), *Kriminalität, Ökonomie und Europäischer Sozialstaat*, Springer, Berlin, pp. 167–185.

Antunes, G. and A. L. Hunt (1973a). "The Deterrent Impact Of Criminal Sanctions: Some Implications For Criminal Justice Policy." *Journal Of Urban Law.* 51 (2), 145–161.

——— and A. L. Hunt (1973b). "The Impact Of Certainty And Severity Of Punishment On Levels Of Crime In American States: An Extended Analysis." *Journal Of Criminal Law And Criminology.* 64 (4), 486–493.

Archambeault, W. G. (1979). "Evaluating The Utility Of The Deterrent Residual In Predicting Deterrent Outcomes In A Self-Report Study: A Comparative Analysis Of The Beccarian And Benthamite Modells Of Deterrence Theory." PhD dissertation.

Avio, K. L. (1979). "Capital Punishment In Canada: A Time-Series Analysis Of The Deterrent Hypothesis." *Canadian Journal Of Economics-Revue Canadienne De Economique.* 12 (4), 647–676.

——— (1987). "Clemency In Economic And Retributive Models Of Punishment." *International Review Of Law And Economics.* 7 (1), 79–88.

——— (1988). "Measurement Errors And Capital Punishment." *Applied Economics.* 20 (9), 1253–1262.

——— and S. C. Clark (1976). *Property Crime In Canada: An Econometric Study*, Ontario Council Economic Research Studies, Toronto.

——— and S. C. Clark (1978). "The Supply Of Property Offences In Ontario: Evidence On The Deterrent Effect Of Punishment." *Canadian Journal Of Economics-Revue Canadienne De Economique.* 11 (1), 1–19.

Ayres, I. and S. D. Levitt (1998). "Measuring Positive Externalities From Unobservable Victim Precaution: An Empirical Analysis Of Lojack." *Quarterly Journal Of Economics.* 113 (1), 43–77.

Bacon, P. and M. O'Donohue (1975). "The Economics Of Crime In The Republic Of Ireland: An Exploratory Paper." *Economic And Social Review.* 7 (1), 19–34.

Bailey, W. C. (1971). "Models Of Deterrence." PhD dissertation, Washington State University, Pullman.

——— (1974). "Murder And The Death Penalty." *Journal Of Criminal Law And Criminology.* 65 (2), 416–423.

——— (1975). "Murder And Capital Punishment: Some Further Evidence." *American Journal Of Orthopsychiatry.* 45 (4), 669–688.

——— (1976). "Use Of Death Penalty Versus Outrage At Murder: Some Additional Evidence And Considerations." *Crime And Delinquency.* 22 (1), 31–39.

——— (1977). "Deterrence And The Violent Sex Offender: Imprisonment Versus The Death Penalty." *Journal Of Behavioral Economics.* 6 (1), 107–145.

——— (1978a). "Deterrence And The Celerity Of The Death Penalty: A Neglected Question In Deterrence Research." Discussion Paper No. 532–78, Institute For Research On Poverty, Wisconsin University.

——— (1978b). "The Deterrent Effect Of Arrest And Conviction On Crime Rates In Indian States And Territories." *Annales Internationales De Criminologie.* 17 (1–2), 21–50.

——— (1978c). "Some Further Evidence On Imprisonment Vs. Death Penalty As Deterrent To Murder." *Law And Human Behavior.* 2 (3), 245–260.

——— (1980a). "Deterrence And The Celerity Of The Death Penalty: A Neglected Question In Deterrence Research." *Social Forces.* 58 (4), 1308–1333.

——— (1980b). "A Multivariate Cross-Sectional Analysis Of The Deterrent Effect Of The Death Penalty." *Sociology And Social Research.* 64 (2), 183–207.

——— (1982). "Capital Punishment And Lethal Assaults Against Police." *Criminology.* 19 (4), 608–625.

——— (1983). "Disaggregation In Deterrence And Death Penalty Research: The Case Of Murder In Chicago." *Journal Of Criminal Law And Criminology.* 74 (3), 827–859.

——— (1990). "Murder, Capital Punishment, And Television: Execution, Publicity, And Homicide Rates." *American Sociological Review.* 55 (5), 628–633.

——— (1991). "The General Prevention Effect Of Capital Punishment For Non-Capital Felonies." in R. M. Bohm (ed.), *The Death Penalty In America: Current Research*, Anderson Publishing Co., Cincinnati, pp. 21–38.

——— (1998). "Deterrence, Brutalization And The Death Penalty: Another Examination Of Oklahoma's Return To Capital Punishment." *Criminology.* 36 (4), 711–733.

——— and R. D. Peterson (1989). "Murder And Capital Punishment: A Monthly Time-Series Analysis Of Execution Publicity." *American Sociological Review.* 54 (5), 722–743.

——— and R. D. Peterson (1994). "Murder, Capital Punishment, And Deterrence: A Review Of The Evidence And An Examination Of Police Killings." *Journal Of Social Issues.* 50 (2), 53–74.

——— and R. P. Lott (1976). "Crime, Punishment And Personality: An Examination Of Deterrence Question." *Journal Of Criminal Law And Criminology.* 67 (1), 99–109.

———, D. J. Martin, and L. N. Gray (1974). "Crime And Deterrence: A Correlation Analysis." *Journal Of Research In Crime And Delinquency.* 11 (2), 124–143.

Baker, T., A. Harel, and T. Kugler (2004). "The Virtues Of Uncertainty In Law: An Experimental Approach." *Iowa Law Review.* 89 (2), 443–494.

Balbo, M. and J. Posadas (1998). "Una Primera Aproximación Al Estudio Del Crimen En La Argentina." Working Paper 1296, Asociacion Argentina De Economia Politica, Buenos Aires.

Bandura, A. (1969). *Principles Of Behavior Modification*, Holt, Rineheart & Winston Inc., New York.

Bar-Ilan, A. and B. Sacerdote (2001). "The Response To Fines And Probability Of Detection In A Series Of Experiments." Working Paper 8638, National Bureau of Economic Research, Cambridge, Massachusetts.

——— and B. Sacerdote (2004). "The Response Of Criminals And Noncriminals To Fines." *Journal Of Law And Economics.* 47 (1), 1–18.

Baron, S. W. and L. W. Kennedy (1998). "Deterrence And Homeless Male Street Youths." *Canadian Journal Of Criminology-Revue Canadienne De Criminologie.* 40 (1), 27–60.

Bartley, W. A., M. A. Cohen, and L. Froeb (1998). "The Effect Of Concealed Weapons Laws: An Extreme Bound Analysis." *Economic Inquiry.* 36 (2), 258–265.

Bean, F. D. and R. G. Cushing (1971). "Criminal Homicide, Punishment, And Deterrence: Methodological And Substantive Reconsiderations." *Social Science Quarterly.* 52 (2), 277–289.

Becker, G. S. (1968). "Crime And Punishment: An Economic Approach." *Journal Of Political Economy.* 76 (2), 169–217.

——— (1993). "Nobel Lecture: The Economic Way Of Looking At Behavior." *Journal Of Political Economy.* 101 (3), 385–409.

Becker, W., H.-J. Büchner, and S. Sleeking (1987). "The Impact Of Public Transfer Expenditures On Tax Evasion: An Experimental Approach." *Journal Of Public Economics.* 34 (2), 234–252.

Becsi, Z. (1999). "Economics And Crime In The States." *Economic Review Of The Federal Reserve Bank Of Atlanta.* 84 (1), 38–56.

Bedard, K. and E. Helland (2000). "The Location Of Women's Prisons And The Deterrence Effect Of 'Harder' Time." Working Paper 2000-06, Claremont College, Claremont.

Benjamini, J. and S. Maital (1985). "Optimal Tax Evasion And Optimal Tax Evasion Policy: Behavioral Aspects." in A. Wenig and W. Gaertner (eds.), *The Economics Of The Shadow Economy*, Vol. 15, Springer, Berlin, pp. 245–264.

Benson, B. L. and B. D. Mast (2001). "Privately Produced General Deterrence." *Journal Of Law And Economics.* 44 (2), 725–746.

———, B. D. Mast, and D. W. Rasmussen (2000). "Can Police Deter Drunk Driving?" *Applied Economics.* 32 (3), 357–366.

——— , D. W. Rasmussen, and B. D. Mast (1999). "Deterring Drunk Driving Fatalities: An Economics Of Crime Perspective." *International Review Of Law And Economics.* 19 (2), 205–225.

——— , D. W. Rasmussen, and I. Kim (1998). "Deterrence And Public Policy: Trade-Offs In The Allocation Of Police Resources." *International Review Of Law And Economics.* 18 (1), 77–100.

——— , I. Kim, D. W. Rasmussen, and T. W. Zuehlke (1992). "Is Property Crime Caused By Drug Use Or By Drug Enforcement Policy?" *Applied Economics.* 24 (7), 679–692.

——— , I. S. Leburn, and D. W. Rasmussen (2001). "The Impact Of Drug Enforcement On Crime: An Investigation Of The Opportunity Cost Of Police Resources." *Journal Of Drug Issues.* 31 (4), 989–1006.

Bentham, J. (1830). *The Rationale Of Punishment*, Nicklin, Philip H., Philadelphia. Written In The 1770s.

Berger, D. E. and J. R. Snortum (1986). "A Structural Model Of Drinking And Driving: Alcohol Consumption, Social Norms, And Moral Commitments." *Criminology.* 24 (1), 139–153.

——— and W. D. Marelich (1997). "Legal And Social Control Of Alcohol-Impaired Driving In California: 1983–1994." *Journal Of Studies On Alcohol.* 58 (5), 518–523.

Berk, R. A. (2005). "New Claims About Executions And General Deterrence: Déjà Vu All Over Again?" *Journal Of Empirical Legal Studies.* 2 (2), 303–330.

Berlitz, C., H.-W. Guth, R. Kaulitzki, and K. F. Schumann (1987). "Grenzen der Generalprävention: Das Beispiel Jugendkriminalität." *Kriminologisches Journal.* 19 (1), 13–31.

Beron, K. J., H. Tauchen, and A. D. Witte (1988). "A Structural Equation Model For Tax Compliance And Auditing." Working Paper 2556, National Bureau of Economic Research, Cambridge, Massachusetts.

Beyleveld, D. (1980). *A Bibliography On General Deterrence*, Saxon House, Teakfield Limited, Westmead.

Bijmolt, T. H. A. and R. G. M. Pieters (2001). "Meta-Analysis In Marketing When Studies Contain Multiple Measurements." *Marketing Letters.* 12 (2), 157–169.

Biles, D. (1982). "Crime And Imprisonment: An Australian Time Series Analysis." *Australian And New Zealand Journal Of Criminology.* 15 (3), 133–153.

Bishop, D. M. (1983). "Deterrence And Social Control: A Longitudinal Study Of The Effects Of Sanctioning And Social Bonding On The Prevention Of Delinquency." PhD dissertation, School Of Criminal Justice, University At Albany.

——— (1984a). "Deterrence: A Panel Analysis." *Justice Quarterly.* 1 (3), 311–328.

——— (1984b). "Legal And Extralegal Barriers To Delinquency: A Panel Analysis." *Criminology.* 22 (3), 403–419.

Black, D. A. and D. S. Nagin (1998). "Do Right-To-Carry Laws Deter Violent Crime?" *Journal Of Legal Studies.* 27 (1), 209–219.

Black, T. and T. Orsagh (1978). “New Evidence On The Efficacy Of Sanctions As A Deterrent To Homicide.” *Social Science Quarterly.* 58 (4), 616–631.

Blau, J. R. and P. M. Blau (1982). “The Cost Of Inequality: Metropolitan Structure And Violent Crime.” *American Sociological Review.* 47 (1), 114–129.

Blau, P. M. and R. M. Golden (1986). “Metropolitan Structure And Criminal Violence.” *Sociological Quarterly.* 27 (1), 15–26.

Blaug, M. (1980). *The Methodology Of Economics*, Cambridge University Press, Cambridge.

Block, M. K. and J. M. Heineke (1975). “A Labor Theoretic Analysis Of The Criminal Choice.” *American Economic Review.* 65 (3), 314–325.

Blumstein, A. and D. S. Nagin (1977). “The Deterrent Effect Of Legal Sanctions On Draft Evasion.” *Stanford Law Review.* 29 (3), 241–276.

Bodman, P. M. and C. Maultby (1997). “Crime, Punishment And Deterrence In Australia: A Further Empirical Investigation.” *International Journal Of Social Economics.* 24 (7), 884–901.

Bohnet, I. and R. D. Cooter (2001). “Expressive Law: Framing Or Equilibrium Selection?” Working Paper No. 1058, Berkeley Olin Program In Law & Economics, University Of California, Berkeley.

Bonger, W. A. (1916). *Criminality And Economic Conditions*, Little, Brown, And Company, Boston.

Bönitz, D. (1989). *Strafgesetze und Verhaltenssteuerung. Zur generalpräventiven Wirksamkeit staatlicher Strafdrohung*, Schwartz, Göttingen.

Borack, J. I. (1998). “An Estimate Of The Impact Of Drug Testing On The Deterrence Of Drug Use.” *Military Psychology.* 10 (1), 17–25.

Bosco, L. and L. Mittone (1997). “Tax Evasion And Moral Constraints: Some Experimental Evidence.” *Kyklos.* 50 (3), 297–324.

Bowers, W. J. and G. L. Pierce (1975). “The Illusion Of Deterrence In Isaac Ehrlich’s Research On Capital Punishment.” *Yale Law Journal.* 85 (2), 187–208.

——— and R. G. Salem (1972). “Severity Of Formal Sanctions As A Repressive Response To Deviant Behavior.” *Law And Society Review.* 6 (3), 427–441.

Braga, A. A., D. M. Kennedy, E. J. Waring, and A. M. Piehl (2001). “Problem-Oriented Policing, Deterrence, And Youth Violence: An Evaluation Of Boston’s Operation Ceasefire.” *Journal Of Research In Crime And Delinquency.* 38 (3), 195–225.

Braithwaite, J. B. (1989). *Crime, Shame And Reintegration*, Cambridge University Press, New York.

——— and T. Makkai (1991). “Testing An Expected Utility Model Of Corporate Deterrence.” *Law And Society Review.* 25 (1), 7–40.

Brezina, T. and A. R. Piquero (2003). “Exploring The Relationship Between Social And Non-Social Reinforcement In The Context Of Social Learning Theory.” in R. Akers and G. F. Jensen (eds.), *Social Learning Theory And The Explanation Of Crime*, New York, pp. 265–288.

Brier, S. S. and S. E. Fienberg (1980). “Recent Econometric Modeling Of Crime And Punishment. Support For The Deterrence Hypothesis?” *Evaluation Review.* 4 (2), 147–191.

Bronars, S. G. and J. R. Lott (1998). “Criminal Deterrence, Geographic Spillovers, And The Right To Carry Concealed Handguns.” *American Economic Review.* 88 (2), 475–479.

Brown, D. W. (1978). “Arrest Rates And Crime Rates: When Does A Tipping Effect Occur?” *Social Forces.* 57 (2), 671–682.

Brumm, H. J. and D. O. Cloninger (1995a). “The Drug War And The Homicide Rate: A Direct Correlation?” *Cato Journal.* 14 (3), 509–517.

——— and D. O. Cloninger (1995b). “Violent Crime And Punishment: An Application Of The Lisrel Model.” *Applied Economics.* 27 (8), 719–725.

——— and D. O. Cloninger (1996). “Perceived Risk Of Punishment And The Commission Of Homicides: A Covariance Structure Analysis.” *Journal Of Economic Behavior And Organization.* 31 (1), 1–11.

Bryjak, G. J. (1980). “Deterrence Theory And Anomie.” PhD dissertation, University of Oklahoma, Norman.

Buikhuisen, W. (1974). “General Deterrence: Research And Theory.” *Abstracts On Criminology And Penology.* 14 (3), 285–298.

Bunn, D. N., S. B. Caudill, and D. M. Gropper (1992). “Crime In The Classroom: An Economic Analysis Of Undergraduate Student Cheating Behavior.” *Journal Of Economic Education.* 23 (3), 197–207.

Buonanno, P. (2003). “Identifying The Effect Of Education On Crime. Evidence From The Italian Regions.” Working Paper 65, Department Of Economics, Università Milano Bicocca, Milano.

Burkett, S. R. and D. A. Ward (1993). “A Note On Perceptual Deterrence, Religiously Based Moral Condemnation, And Social Control.” *Criminology.* 31 (1), 119–134.

——— and E. L. Jensen (1975). “Conventional Ties, Peer Influence, And The Fear Of Apprehension: A Study Of Adolescent Marijuana Use.” *Sociological Quarterly.* 16 (4), 522–533.

——— and W. T. Carrithers (1980). “Adolescents‘ Drinking And Perceptions Of Legal And Informal Sanctions: A Test Of Four Hypotheses.” *Journal Of Studies On Alcohol.* 41 (9), 839–853.

Burnell, J. D. (1988). “Crime And Racial Composition In Contiguous Communities As Negative Externalities: Prejudiced Households’ Evaluation Of Crime Rate And Segregation Nearby Reduces Housing Values And Tax Revenues.” *American Journal Of Economics And Sociology.* 47 (2), 177–193.

Bursik, R. J. and Y. Baba (1986). “Individual Variations In Crime-Related Decisions.” *Social Science Research.* 15 (1), 71–81.

———, H. G. Grasmick, and M. B. Chamlin (1990). “The Effect Of Longitudinal Arrest Patterns On The Development Of Robbery Trends At The Neighborhood Level.” *Criminology.* 28 (3), 431–450.

Cameron, S. (1988). “The Economics Of Crime Deterrence: A Survey Of Theory And Evidence.” *Kyklos.* 41 (2), 301–323.

——— (1989). “Victim Compensation Does Not Increase The Supply Of Crime.” *Journal Of Economic Studies.* 16 (4), 52–59.

——— (1991). “Policing In The Uneconomic Zone Of The Production Function.” *Journal Of Socio-Economics.* 20 (4), 313–323.

Campbell, D. T. and L. H. Ross (1968). “The Connecticut Crackdown On Speeding: Time-Series Data In Quasi Experimental Analysis.” *Law And Society Review.* 3 (1), 33–53.

Cappell, C. and G. Sykes (1991). “Prison Commitments, Crime, And Unemployment: A Theoretical And Empirical Specification For The United States, 1933–1985.” *Journal Of Quantitative Criminology.* 7 (2), 155–199.

Cardarelli, A. P. (1968). “An Analysis Of Police Killed By Criminal Action: 1961–1963.” *Journal Of Criminal Law, Criminology And Police Science.* 59 (3), 447–453.

Carmichael, F. and R. Ward (2000). “Youth Unemployment And Crime In The English Regions And Wales.” *Applied Economics.* 32 (5), 559–571.

Carmichael, S. and A. R. Piquero (2004). “Sanctions, Perceived Anger, And Criminal Offending.” *Journal Of Quantitative Criminology.* 20 (4), 371–393.

———, L. Langton, G. Pendell, J. D. Reitzel, and A. R. Piquero (2005). “Do The Experiential And Deterrent Effect Operate Differently Across Gender?” *Journal Of Criminal Justice.* 33 (3), 267–276.

Carr-Hill, R. A. and N. H. Stern (1973). “An Econometric Model Of The Supply And Control Of Recorded Offences In England And Wales.” *Journal Of Public Economics.* 2 (4), 289–318.

Casey, J. T. and J. T. Scholz (1991). “Beyond Deterrence: Behavioral Decision Theory And Tax Compliance.” *Law And Society Review.* 25 (4), 821–844.

Caudill, B. D., G. K. Kantor, and S. Ungerleider (1990). “Driving While Intoxicated: Increased Deterrence Or Alternative Transportation For The Drunk Driver.” *Journal Of Substance Abuse.* 2 (1), 51–67.

Cerro, A. M. and O. Meloni (2000). “Determinants Of The Crime Rate In Argentina During The ‘90s.” *Estudios De Economia.* 27 (2), 297–311.

Chadwick, E. (1829). “Preventive Police.” *London Review.* 1, 252–308.

Chaiken, J. M., M. W. Lawless, and K. A. Stevenson (1974). *The Impact Of Police Activity On Subway Crime*, Rand Corp., Santa Monica.

Chaloupka, F. J., H. Saffer, and M. Grossman (1993). “Alcohol Control Policies And Motor Vehicle Fatalities.” *Journal Of Legal Studies.* 22 (1), 161–186.

Chambers, L., R. Roberts, and C. Voelker (1976). “The Epidemiology Of Traffic Accidents And The Effect Of The 1969 Breathalyser Amendment In Canada.” *Accident Analysis And Prevention.* 8 (2), 201–206.

Chambouleyron, A. and M. Willington (1998). “Crime And Punishment In Argentina: An Empirical Approach.” Working Paper 1309, Asociacion Argentina De Economia Politica, Buenos Aires.

Chamlin, M. B. and J. K. Cochran (1997). “Social Altruism And Crime.” *Criminology.* 35 (2), 203–227.

——— , H. G. Grasmick, and R. J. Bursik (1992). “Time Aggregation And Time Lag In Macro-Level Deterrence Research.” *Criminology.* 30 (3), 377–396.

Chapman, J. I. (1976). “An Economic Model Of Crime And Police: Some Empirical Results.” *Journal Of Research In Crime And Delinquency.* 13 (1), 48–63.

Cheatwood, D. (1993). “Capital Punishment And The Deterrence Of Violent Crime In Comparable Counties.” *Criminal Justice Review.* 18 (2), 165–181.

Cherry, T. L. (1999). “Unobserved Heterogeneity Bias When Estimating The Economic Model Of Crime.” *Applied Economics Letters.* 6 (11), 753–757.

——— (2001). “Financial Penalties As An Alternative Criminal Sanction: Evidence From Panel Data.” *Atlantic Economic Journal.* 29 (4), 450–458.

——— and J. A. List (2002). “Aggregation Bias In The Economic Model Of Crime.” *Economics Letters.* 75 (1), 81–86.

Chilton, R. (1982). “Analyzing Urban Crime Data: Deterrence And The Limitations Of Arrests Per Offense Ratios.” *Criminology.* 19 (4), 590–607.

Chipman, H., E. I. George, and R. E. McCulloch (2001). “The Practical Implementation Of Bayesian Model Selection.” in P. Lahiri (ed.), *Model Selection*, Vol. 38 of *Lecture Notes - Monograph Series*, Beachwood, pp. 65–116.

Chiricos, T. G. and G. P. Waldo (1970). “Punishment And Crime: An Examination Of Some Empirical Evidence.” *Social Problems.* 18 (2), 200–217.

Cho, Y. H. (1972). “A Multiple Regression Model For The Measurement Of The Public Policy Impact On Big City Crime.” *Policy Sciences.* 3 (4), 435–455.

Chressanthis, G. A. (1989). “Capital Punishment And The Deterrent Effect Revisited: Recent Time-Series Econometric Evidence.” *Journal Of Behavioral Economics.* 18 (2), 81–97.

——— and P. W. Grimes (1990). “Criminal Behaviour And Youth In The Labour Market: The Case Of The Pernicious Minimum Wage.” *Applied Economics.* 22 (11), 1495–1508.

Clark, A. A. (1984). “Employee Theft And Methods Of Deterrence.” PhD dissertation, Texas Woman’s University, Denton.

Clark, D. E. and J. C. Cosgrove (1990). “Hedonic Prices, Identification, And The Demand For Public Safety.” *Journal Of Regional Science.* 30 (1), 105–121.

Clarke, R. V. (1966). “Approved School Boy Absconders And Corporal Punishment.” *British Journal Of Criminology.* 6 (4), 364–375.

Clements, M. P. and D. I. Harvey (2004). “Forecast Encompassing Tests And Probability Forecasts.” Discussion Paper, School Of Economics, University Of Nottingham, Nottingham.

Cloninger, D. O. (1975). “The Deterrence Effect Of Law Enforcement: An Evaluation Of Recent Findings And Some New Evidence.” *American Journal Of Economics And Sociology.* 34 (3), 323–335.

——— (1977). “Deterrence And The Death Penalty: A Cross Sectional Analysis.” *Journal Of Behavioral Economics.* 6 (1), 87–107.

——— (1981). “Risk, Arson And Abandonment.” *Journal Of Risk And Insurance.* 48 (3), 494–504.

——— (1987). “Capital Punishment And Deterrence: A Revision.” *Journal Of Behavioral Economics.* 16 (4), 55–58.

——— (1990). “Arson And Abandonment: A Restatement.” *Journal Of Risk And Insurance.* 57 (3), 540–545.

——— (1992). “Capital Punishment And Deterrence: A Portfolio Approach.” *Applied Economics.* 24 (6), 635–645.

——— (1994). “Enforcement Risk And Deterrence: A Re-Examination.” *Journal Of Socio-Economics.* 23 (3), 273–285.

——— and R. Marchesini (2001). “Execution And Deterrence: A Quasi-Controlled Group Experiment.” *Applied Economics.* 33 (5), 569–576.

Clotfelter, C. T. (1978). “Private Security And The Public Safety.” *Journal Of Urban Economics.* 5 (3), 388–402.

Coakley, J., A.-M. Fuertes, and R. Smith (2006). “Unobserved Heterogeneity In Panel Time Series Models.” *Computational Statistics And Data Analysis.* 9 (1), 2361–2380.

Cochran, J. K. and M. B. Chamlin (2000). “Deterrence And Brutalization: The Dual Effects Of Executions.” *Justice Quarterly.* 17 (4), 685–706.

——— , M. B. Chamlin, and M. Seth (1994). “Deterrence Or Brutalization? An Impact Assessment Of Oklahoma’s Return To Capital Punishment.” *Criminology.* 32 (1), 107–134.

——— , M. B. Chamlin, P. B. Wood, and C. S. Sellers (1999). “Shame, Embarrassment, And Formal Sanction Threats: Extending The Deterrence Rational Choice Model To Academic Dishonesty.” *Sociological Inquiry.* 69 (1), 91–105.

Cohen, L. (1978). “Sanction Threats And Violation Behavior: An Inquiry Into Perceptual Variation.” in C. Wellford (ed.), *Quantitative Studies In Criminology*, Beverly Hills, pp. 84–99.

Cohen, L. E. and K. C. Land (1984). “Discrepancies Between Crime Reports And Crime Surveys: Urban And Structural Determinants.” *Criminology.* 22 (4), 499–530.

——— and K. C. Land (1987). “Age Structure And Crime: Symmetry Versus Asymmetry And The Projection Of Crime Rates Through The 1990s.” *American Sociological Review.* 52 (2), 170–183.

——— and M. Felson (1979). “Social Change And Crime Rate Trends: A Routine Activity Approach.” *American Sociological Review.* 44 (4), 588–608.

——— , M. Felson, and K. C. Land (1980). “Property Crime Rates In The United States: A Macrodynamic Analysis, 1947–1977; With Ex Ante Forecasts For The Mid-1980s.” *American Journal Of Sociology.* 86 (1), 90–118.

Cohen, M. A. (1987). “Optimal Enforcement Strategy To Prevent Oil Spills: An Application Of A Principal-Agent Model With Moral Hazard.” *Journal Of Law And Economics.* 30 (1), 23–51.

——— (2000). “Measuring The Costs And Benefits Of Crime And Justice.” in D. Duffee (ed.), *Measurement And Analysis Of Crime And Justice*, Vol. 4 of *Criminal Justice*, National Institute Of Justice, Washington D.C., pp. 263–315.

——— and S. S. Simpson (1997). “Corporate Criminal Liability.” in W. S. Lofquist, M. A. Cohen, and G. Rabe (eds.), *Debating Corporate Crime*, Anderson Publishing And Academy Of Criminal Justice Sciences, Cincinatti. Chapter 2.

Collins, J. H. and R. D. Plumlee (1991). “The Taxpayer’s Labor And Reporting Decision: The Effect Of Audit Schemes.” *Accounting Review.* 66 (3), 559–576.

Cook, P. J. (1977). “Punishment And Crime: A Critique Of Current Findings Concerning The Preventive Effects Of Punishment.” *Law And Contemporary Problems.* 41 (1), 164–204.

——— and J. Ludwig (2002). “The Effects Of Gun Prevalence On Burglary: Deterrence Vs Inducement.” Working Paper 8926, National Bureau of Economic Research, Cambridge, Massachusetts.

Corman, H. (1981). “Criminal Deterrence In New York: The Relationship Between Court Activities And Crime.” *Economic Inquiry.* 19 (3), 476–487.

——— and N. H. Mocan (2000). “A Time-Series Analysis Of Crime, Deterrence, And Drug Abuse In New York City.” *American Economic Review.* 90 (3), 584–604.

——— and N. H. Mocan (2002). “Carrots, Sticks And Broken Windows.” Working Paper 9061, National Bureau of Economic Research, Cambridge, Massachusetts.

——— and T. Joyce (1990). “Urban Crime Control: Violent Crimes In New York City.” *Social Science Quarterly.* 71 (3), 567–584.

——— , T. Joyce, and N. Lovitch (1987). “Crime, Deterrence And The Business Cycle In New York City: A VAR Approach.” *Review Of Economics And Statistics.* 69 (4), 695–700.

Cornwell, C. and W. N. Trumbull (1994). “Estimating The Economic Model Of Crime With Panel Data.” *Review Of Economics And Statistics.* 76 (2), 360–366.

Cover, J. P. and P. D. Thistle (1988). “Time Series, Homicide, And The Deterrent Effect Of Capital Punishment.” *Southern Economic Journal.* 54 (3), 615–622.

Craig, S. G. and E. J. Heikkila (1989). “Urban Safety In Vancouver: Allocation And Production Of A Congestible Public Good.” *Canadian Journal Of Economics.* 22 (4), 867–884.

Cullen, F. T. (1994). “Social Support As An Organizing Concept For Criminology: Presidential Address To The Academy Of Criminal Justice Sciences.” *Justice Quarterly.* 11 (4), 527–559.

Cummings, R. G., J. Martinez-Vazquez, and M. McKee (2001). “Cross Cultural Comparisons Of Tax Compliance Behavior.” Working Paper 01–03, Andrew Young School Of Policy Studies, Georgia State University.

Curti, H. (1984). “Zur Abschreckungswirkung Strafrechtlicher Sanktionen In Der Bundesrepublik Deutschland: Eine Empirische Untersuchung.” in C. Ott and H.-B. Schäfer (eds.), *Die Präventivwirkung zivil- und strafrechtlicher Sanktionen*, Mohr-Siebeck, Tübingen, pp. 71–94.

——— (1999). *Abschreckung durch Strafe: Eine ökonomische Analyse der Kriminalität*, Deutscher Universitätsverlag, Wiesbaden.

Danziger, S. and D. Wheeler (1975). “The Economics Of Crime: Punishment Or Income Redistribution.” *Review Of Social Economy.* 33 (2), 113–131.

de Long, J. B. and K. Lang (1992). “Are All Economic Hypotheses False?” *Journal Of Political Economy.* 100 (6), 1257–1272.

Deadman, D. F. (2003). “Forecasting Residential Burglary.” *International Journal Of Forecasting.* 19 (4), 567–578.

——— and D. J. Pyle (2000). “An Economic Model Of Criminal Activity.” in Z. McDonald and D. Pyle (eds.), *Illicit Activity: The Economics Of Crime, Drugs And Tax Fraud*, Ashgate Publishers, Aldershot, pp. 15–38.

Decker, S. H. (1976). “Criminalization, Victimization And Structural Correlates Of Twenty-Six American Cities.” PhD dissertation, Florida State University, Tallahassee.

——— and C. W. Kohfeld (1984). “A Deterrence Study Of The Death Penalty In Illinois, 1933–1980.” *Journal Of Criminal Justice.* 12 (4), 367–377.

——— and C. W. Kohfeld (1990a). “Certainty, Severity, And The Probability Of Crime: A Logistic Analysis.” *Policy Studies Journal.* 19 (1), 2–21.

——— and C. W. Kohfeld (1990b). “The Deterrent Effect Of Capital Punishment In The Five Most Active Execution States: A Time Series Analysis.” *Criminal Justice Review.* 15 (2), 173–191.

DeFina, R. H. and T. M. Arvanites (2002). “The Weak Effect Of Imprisonment On Crime: 1971–1998.” *Social Science Quarterly.* 83 (3), 635–653.

Demers, D. K. and R. J. Lundman (1987). “Perceptual Deterrence Research: Some Additional Evidence For Designing Studies.” *Journal Of Quantitative Criminology.* 3 (2), 185–194.

Denny, K., C. Harmon, R. Lydon, and I. Walker (2004). “An Econometric Analysis Of Burglary In Ireland.” Discussion Paper 2004–04, Institute For The Study Of Social Change, University College Dublin, Dublin.

Denq, F., M. S. Vaughn, and F. F. Huang (1994). “Correlates Of Crime In Taiwan: A Time-Series Analysis From 1964 To 1990.” *Crime Law And Social Change.* 21 (3), 267–285.

Deshapriya, E. B. and N. Iwase (1996). “Are Lower Legal Blood Alcohol Limits And A Combination Of Sanctions Desirable In Reducing Drunken Driver-Involved Traffic Fatalities And Traffic Accidents?” *Accident Analysis And Prevention.* 28 (6), 721–731.

DeSimone, J. (2001). “The Effect Of Cocaine Prices On Crime.” *Economic Inquiry.* 39 (4), 627–643.

——— and M. C. Farrelly (2003). “Price And Enforcement Effects On Cocaine And Marijuana Demand.” *Economic Inquiry.* 41 (1), 98–115.

Devine, J. A., J. F. Sheley, and D. M. Smith (1988). “Macroeconomic And Social-Control Policy Influences On Crime Rate Changes, 1948–1985.” *American Sociological Review.* 53 (3), 407–420.

Dezhbakhsh, H. and J. M. Shepherd (2003). “The Deterrent Effect Of Capital Punishment: Evidence From A Judicial Experiment.” Working Paper 03–14, Department Of Economics, Emory University, Atlanta.

——— and P. H. Rubin (1998). “Lives Saved Or Lives Lost? The Effects Of Concealed-Handgun Laws On Crime.” *American Economic Review.* 88 (2), 468–474.

——— , P. H. Rubin, and J. M. Shepherd (2001). “Does Capital Punishment Have A Deterrent Effect? New Evidence From Post-Moratorium Panel Data.” *American Law And Economics Review.* 5 (2), 344–376.

Diez-Ticio, A. (2000). “The Relationship Between Economic Conditions And Property Crime: Evidence For The United States.” in Z. McDonald and D. Pyle (eds.), *Illicit Activity: The Economics Of Crime, Drugs And Tax Fraud*, Ashgate Publishers, Aldershot, pp. 103–128.

Dijkstra, T. K. (1995). “Pyrrho’s Lemma, Or Have It Your Way.” *Metrika.* 42 (1), 119–125.

DiPasquale, D. and E. L. Glaeser (1998). “The Los Angeles Riot And The Economics Of Urban Unrest.” *Journal Of Urban Economics.* 43 (1), 52–78.

Dittrich, M. and G. Markwardt (2004). “Arbeitslosigkeit und Kriminalität: Eine mögliche doppelte Dividende der Arbeitsmarktpolitik?” *ifo Dresden berichtet.* 6, 11–17.

Dölling, D. (1983). “Strafeinschätzung und Delinquenz bei Jugendlichen und Heranwachsenden: Ein Beitrag zur Empirischen Analyse der generalpräventiven Wirkungen der Strafe.” in H.-J. Kerner, H. Kury, and K. Sessar (eds.), *Deutsche Forschungen zur Kriminalitätsentstehung und Kriminalitätskontrolle*, Heymanns Verlag, Köln, pp. 51–85.

——— , D. Hermann, H. Entorf, A. Woll, and T. Rupp (2007). “Metaanalyse empirischer Abschreckungsstudien: Untersuchungsansatz und erste empirische Befunde.” in F. Lösel, D. Bender, and J.-M. Jehle (eds.), *Kriminologie und wissensbasierte Kriminalpolitik: Entwicklungs- und Evaluationsforschung*, in ‘Neue kriminologische Schriftenreihe.’, Forum Verlag, Godesberg, pp. 632–648.

——— , H. Entorf, D. Hermann, T. Rupp, A. Woll, and A. Häring (2006). “Zur Generalpräventiven Abschreckungswirkung des Strafrechts - Befunde einer Metaanalyse.” *Soziale Probleme.* 17 (2), 193–209.

Donohue, J. J. and J. Wolfers (2005). “Uses And Abuses Of Empirical Evidence In The Death Penalty Debate.” *Stanford Law Review.* 58 (3), 791–845.

——— and S. D. Levitt (2000). “The Impact Of Legalized Abortion On Crime.” Working Paper 8004, National Bureau of Economic Research, Cambridge, Massachusetts.

Doyle, J. M., E. Ahmed, and R. N. Horn (1999). "The Effects Of Labor Markets And Income Inequality On Crime: Evidence From Panel Data." *Southern Economic Journal.* 65 (4), 717–738.

Dubin, J. A. and L. L. Wilde (1988). "An Empirical Analysis Of Federal Income Tax Auditing And Compliance." *National Tax Journal.* 41 (1), 61–74.

Dugan, L., G. LaFree, and A. R. Piquero (2005). "Testing A Rational Choice Model Of Airline Hijackings." *Criminology.* 43 (4), 1031–1065.

Earnhart, D. (2004a). "Panel Data Analysis Of Regulatory Factors Shaping Environmental Performance." *Review Of Economics And Statistics.* 86 (1), 391–401.

——— (2004b). "Regulatory Factors Shaping Environmental Performance At Publicly Owned Treatment Plants." *Journal Of Environmental Economics And Management.* 48 (1), 655–681.

Edwards, G. C. (1980). *Implementing Public Policy*, Congressional Quarterly Press, Washington D.C.

Ehrlich, I. (1972). "The Deterrent Effect Of Criminal Law Enforcement." *Journal Of Legal Studies.* 1 (2), 259–276.

——— (1973). "Participation In Illegitimate Activities: A Theoretical And Empirical Investigation." *Journal Of Political Economy.* 81 (3), 521–565.

——— (1975a). "The Deterrent Effect Of Capital Punishment: A Question Of Life And Death." *American Economic Review.* 65 (3), 397–417.

——— (1975b). "On The Relation Between Education And Crime." in F. T. Juster (ed.), *Education, Income And Human Behavior*, McGraw-Hill, New York.

——— (1977a). "Capital Punishment And Deterrence: Some Further Thoughts And Additional Evidence." *Journal Of Political Economy.* 85 (4), 741–788.

——— (1977b). "The Deterrent Effect Of Capital Punishment: Reply." *American Economic Review.* 67 (3), 452–458.

——— (1981). "On The Usefulness Of Controlling Individuals: An Economic Analysis Of Rehabilitation, Incapacitation, And Deterrence." *American Economic Review.* 71 (3), 307–322.

——— and G. D. Brower (1987). "On The Issue Of Causality In The Economic Model Of Crime And Law Enforcement: Some Theoretical Considerations And Experimental Evidence." *American Economic Review.* 77 (2), 99–106.

——— and L. Zhiqiang (1999). "Sensitivity Analyses Of The Deterrence Hypothesis: Let's Keep The Econ In Econometrics." *Journal Of Law And Economics.* 42 (1), 455–487.

Eide, E., J. Aasness, and T. Skjerpen (1994). *Economics Of Crime: Deterrence And The Rational Offender.*

Eiser, J. R. (1986). *Social Psychology: Attitudes, Cognition And Behavior*, Cambridge University Press, Cambridge.

Elffers, H., P. van der Heijden, and M. Hezemans (2003). "Explaining Regulatory Non-Compliance: A Survey Study Of Rule Transgression For Two Dutch Instrumental Laws, Applying The Randomized Response Method." *Journal Of Quantitative Criminology.* 19 (4), 409–439.

Elis, L. A. and S. S. Simpson (1995). "Informal Sanction Threats And Corporate Crime: Additive Versus Multiplicative Models." *Journal Of Research In Crime And Delinquency.* 32 (4), 399–424.

Emons, W. (2003). "A Note On The Optimal Punishment For Repeat Offenders." *International Review Of Law And Economics.* 23 (3), 253–259.

Ensor, T. and C. Godfrey (1993). "Modelling The Interactions Between Alcohol, Crime And The Criminal Justice System." *Addiction.* 88 (4), 477–487.

Entorf, H. (1996). "Kriminalität und Ökonomie: Übersicht und neue Evidenz." *Zeitschrift für Wirtschafts- und Sozialwissenschaften.* 116 (3), 417–450.

——— (2003). "Strafe, Haft und Bewährung: Eine Panelökonometrische Analyse der relativen Effizienz Deutscher Strafverfolgungsorgane." Conference Paper.

——— and H. Spengler (2000). "Socio-Economic And Demographic Factors Of Crime In Germany: Evidence From Panel Data Of The German States." *International Review Of Law And Economics.* 20 (1), 75–106.

——— and H. Spengler (2002). *Crime In Europe: Causes And Consequences*, Springer, Berlin.

——— and H. Spengler (2005). "Eine ökonometrische Analyse der Wirkung des Deutschen Strafverfolgungssystems auf das Kriminalitätsaufkommen." Research Note 5, Deutsches Institut für Wirtschaftsforschung, Berlin.

——— and J. Antony (2002). "Zur Gültigkeit der Abschreckung im Sinne der ökonomischen Theorie der Kriminalität - Grundzüge einer Meta-Studie." Discussion Paper 116, Department Of Economics, Darmstadt University Of Technology, Darmstadt.

——— and P. Winker (2002). "Investigating The Drugs-Crime Channel In Economics Of Crime Models: Empirical Evidence From Panel Data Of The German States." Discussion Paper 01–37, Updated, Zentrum für Europäische Wirtschaftsforschung, Mannheim.

Epple, D. and M. Visscher (1984). "Environmental Pollution: Modeling Occurrence, Detection, And Deterrence." *Journal Of Law And Economics.* 27 (1), 29–60.

Erard, B. (1992). "The Influence Of Tax Audits On Reporting Behavior." in J. Slemrod (ed.), *Why People Pay Taxes: Tax Compliance And Enforcement*, University Of Michigan Press, Ann Arbor, pp. 95–115.

Erickson, M. L. and J. P. Gibbs (1973). "The Deterrence Question: Some Alternative Methods Of Analysis." *Social Science Quarterly.* 54 (3), 534–551.

——— and J. P. Gibbs (1975). "Specific Versus General Properties Of Legal Punishments And Deterrence." *Social Science Quarterly.* 56 (3), 390–397.

——— and J. P. Gibbs (1976). "Further Findings On The Deterrence Question And Strategies For Future Research." *Journal Of Criminal Justice.* 4 (3), 175–189.

——— and J. P. Gibbs (1978). "Objective And Perceptual Properties Of Legal Punishment And The Deterrence Doctrine." *Social Problems.* 25 (3), 253–264.

———, J. P. Gibbs, and G. F. Jensen (1977). "The Deterrence Doctrine And The Perceived Certainty Of Legal Punishments." *American Sociological Review.* 42 (2), 305–317.

Erickson, P. G. (1976). "Deterrence And Deviance: The Example Of Cannabis Prohibition." *Journal Of Criminal Law And Criminology.* 67 (2), 222–232.

Estes, W. K. (1944). "An Experimental Study Of Punishment." *Psychological Monographs.* 57 (263), 1–40.

Evans, W. N., D. Neville, and J. D. Graham (1991). "General Deterrence Of Drunk Driving: Evaluation Of Recent American Policies." *Risk Analysis.* 11 (2), 279–289.

Fabrikant, R. (1979). "The Distribution Of Criminal Offenses In An Urban Environment: A Spatial Analysis Of Criminal Spillovers And Of Juvenile Offenders." *American Journal Of Economics And Sociology.* 38 (1), 31–48.

Fajnzylber, P., D. Lederman, and N. Loayza (1998). "Determinants Of Crime Rates In Latin America And The World: An Empirical Assessment." Working Paper, World Bank Latin American And Caribbean Studies, Washington D.C.

———, D. Lederman, and N. Loayza (2000). "Crime And Victimization: An Economic Perspective." *Economia.* 1 (1), 219–302.

———, D. Lederman, and N. Loayza (2002a). "Inequality And Violent Crime." *Journal Of Law And Economics.* 45 (1), 1–40.

———, D. Lederman, and N. Loayza (2002b). "What Causes Violent Crime?" *European Economic Review.* 46 (7), 1323–1357.

Farley, J. E. and M. Hansel (1981). "The Ecological Context Of Urban Crime. A Further Exploration." *Urban Affairs Review (Urban Affairs Quarterly).* 17 (1), 37–54.

Farrelly, M. C., J. W. Bray, G. A. Zarkin, B. W. Wendling, and R. L. Pacula (1999). "The Effects Of Prices And Policies On The Demand For Marijuana: Evidence From The National Household Surveys On Drug Abuse." Working Paper 6940, National Bureau of Economic Research, Cambridge, Massachusetts.

Fattah, E. A. (1983). "A Critique Of Deterrence Research With Particular Reference To The Economic-Approach." *Canadian Journal Of Criminology-Revue Canadienne De Criminologie.* 25 (1), 79–90.

Fehr, E. and S. Gächter (2000). "Cooperation And Punishment In Public Goods Experiments." *American Economic Review.* 90 (4), 980–994.

Feld, L. P. and B. S. Frey (2004). "Illegal, Immoral, Fattening Or What? How Deterrence And Responsive Regulation Shape Tax Morale." Marburg Working Papers on Economics 2004,26.

Fernández, C., E. Ley, and M. F. J. Steel (2001a). "Benchmark Priors For Bayesian Model Averaging." *Journal Of Econometrics.* 100 (2), 381–427.

——— , E. Ley, and M. F. J. Steel (2001b). "Model Uncertainty In Cross-Country Growth Regressions." *Journal Of Applied Econometrics.* 16 (5), 563–576.

Field, S. (1990). "Trends In Crime And Their Interpretation: A Study Of Recorded Crime In Post-War England And Wales." Home Office Research Study 119, Economics And Resources Analysis Unit, Home Office, London.

Fischer, C. S. (1975). "Toward A Subcultural Theory Of Urbanism." *American Journal Of Sociology.* 80 (6), 1319–1341.

Fischer, J. A. (2004). "Determinants Of Crime For Swiss Cantons With Particular Reference To Direct Legislation." Conference Paper.

Fishman, G., S. Hakim, and Y. Shachmurove (1998). "The Use Of Household Survey Data: The Probability Of Property Crime Victimization." *Journal Of Economic And Social Measurement.* 24 (1), 1 –13.

Florax, R. J. G. M. and H. L. F. de Groot (2002). "The Empirical Economic Growth Literature: Robustness, Significance And Size." Discussion Paper 02–040/3, Tinbergen Institute, Amsterdam.

——— , H. L. F. de Groot, and R. A. de Mooij (2002). "Meta-Analysis: A Tool For Upgrading Inputs Of Macroeconomic Policy Models." Discussion Paper 02–041/3, Tinbergen Institute, Amsterdam.

Foege, W. H., M. L. Rosenberg, and J. A. Mercy (1995). "Public Health And Violence Prevention." *Current Issues In Public Health.* 1 (1), 2–9.

Foglia, W. D. (1997). "Perceptual Deterrence And The Mediating Effect Of The Internalized Norms Among Inner-City Teenagers." *Journal Of Research In Crime And Delinquency.* 34 (4), 414–442.

Forst, B. E. (1976). "Participation In Illegitimate Activities: Further Empirical Findings." *Policy Analysis.* 2 (3), 477–492.

——— (1977). "The Deterrent Effect Of Capital Punishment: A Cross-State Analysis Of The 1960's." *Minnesota Law Review.* 61 (5), 743–767.

——— (1983). "Capital Punishment And Deterrence: Conflicting Evidence." *Journal Of Criminal Law And Criminology.* 74 (3), 927–942.

Fortin, B., G. Lacroix, and M.-C. Villeval (2004). "Tax Evasion And Social Interactions." Working Paper 2004S-61, Centre Interuniversitaire De Recherche En Analyse Des Organisations (CIRANO), Montreal.

Foss, R. D., J. S. Richard, and W. R. Donald (1998). "Evaluation Of The Effects Of North Carolina's 0.08% BAC Law." Working Paper, Highway Safety Research Center, University Of North Carolina.

Fox, J. A. (1977). “The Identification And Estimation Of Deterrence: An Evaluation Of Yunker’s Model.” *Journal Of Behavioral Economics.* 6 (1–2), 225–242.

Free Encyclopedia (2008). *Criminal Careers - Bibliography.* http://law.jrank.org/pages/834/criminal-careers-bibliography.html, Accessed 2008.

Freeman, J., P. Liossis, C. Schonfeld, M. Sheehan, V. Siskind, and B. Watson (2006). “The Self-Reported Impact Of Legal And Non-Legal Sanctions On A Group Of Recidivist Drink Drivers.” *Transportation Research Part F: Traffic Psychology And Behaviour.* 9 (1), 53–64.

French, M. T., C. M. Roebuck, and P. K. Alexandre (2004). “To Test Or Not To Test: Do Workplace Drug Testing Programs Discourage Employee Drug Use?” *Social Science Research.* 33 (1), 45–63.

Frey, B. S. and L. P. Feld (2002). “Deterrence And Morale In Taxation: An Empirical Analysis.” Working Paper 760, Center For Economic Studies, München.

Friedland, N. (1985). “Variable Credibility Threats: A Procedure For The Enhancement Of Deterrence Effectiveness.” *Journal Of Applied Social Psychology.* 15 (3), 230–236.

——— , S. Maital, and A. Rutenberg (1978). “A Simulation Study Of Income Tax Evasion.” *Journal Of Public Economics.* 10 (1), 107–116.

Friedman, J., S. Hakim, and U. Spiegel (1989). “The Difference Between Short And Long Run Effects Of Police Outlays On Crime: Policing Deters Criminals Initially, But Later They May Learn By Doing.” *American Journal Of Economics And Sociology.* 48 (2), 177–191.

Friedman, S. R., H. L. F. Cooper, B. Tempalski, M. Keem, R. Friedman, P. L. Flom, and D. C. D. Jarlais (2006). “Relationships Of Deterrence And Law Enforcement To Drug-Related Harms Among Drug Injectors In Us Metropolitan Areas.” *Aids.* 20 (1), 93–99.

Fujii, E. T. and J. Mak (1979). “The Impact Of Alternative Regional Development Strategies On Crime Rates: Tourism Vs. Agriculture In Hawaii.” *Annals Of Regional Science.* 13 (3), 42–56.

——— and J. Mak (1980). “Tourism And Crime: Implications For Regional Development Policy.” *Regional Studies.* 14 (1), 27–36.

Funk, P. and P. Kugler (2003a). “Dynamic Interactions Between Crimes.” *Economics Letters.* 79 (3), 291–298.

——— and P. Kugler (2003b). “Identifying Efficient Crime-Combating Policies By VAR Models: The Example Of Switzerland.” *Contemporary Economic Policy.* 21 (4), 525–538.

Furlong, W. J. (1991). “The Deterrent Effect Of Regulatory Enforcement In The Fishery.” *Land Economics.* 67 (1), 116–129.

——— and S. L. Mehay (1981). “Urban Law Enforcement In Canada: An Empirical Analysis.” *Canadian Journal Of Economics-Revue Canadienne De Economique.* 14 (1), 44–57.

Gabor, T. (1981). “The Crime Displacement Hypothesis: An Empirical Examination.” *Crime And Delinquency.* 27 (3), 390–404.

Gahagan, J., J. T. Tedeschi, T. Faley, and S. Lindskold (1970). “Patterns Of Punishment And Reactions To Threats.” *Journal Of Social Psychology.* 80, 115–116.

Gainey, R. R. and B. K. Payne (2003). "Changing Attitudes Toward House Arrest With Electronic Monitoring: The Impact Of A Single Presentation?" *International Journal Of Offender Therapy And Comparative Criminology.* 47 (2), 196–209.

Garoupa, N. (1999). "Optimal Law Enforcement With Dissemination Of Information." *European Journal Of Law And Economics.* 7 (3), 183–196.

——— and M. Jellal (2002). "Information, Corruption And Optimal Law Enforcement." Discussion Paper 3560, CEPR.

Gartner, R. (1990). "The Victims Of Homicide: A Temporal And Cross-National Comparison." *American Sociological Review.* 55 (1), 92–106.

Gawande, K. and A. K. Bohara (2005). "Agency Problems In Law Enforcement: Theory And Application To The U.S. Coast Guard." *Management Sciences.* 51 (11), 1593–1609.

——— and T. Wheeler (1999). "Measures Of Effectiveness For Governmental Organizations." *Management Sciences.* 45 (1), 42–58.

Geerken, M. and W. R. Gove (1977). "Deterrence, Overload, And Incapacitation: An Empirical Evaluation." *Social Forces.* 56 (2), 424–447.

Gertz, M. G. and L. C. Gould (1995). "Fear Of Punishment And The Willingness To Engage In Criminal Behavior: A Research Note." *Journal Of Criminal Justice.* 23 (4), 377–384.

Gërxhani, K. and A. Schram (2006). "Tax Evasion And Income Source: A Comparative Experimental Study." *Journal of Economic Psychology.* 27 (3), 402–422.

Ghali, M. A., E. Estores, F. Okano, and R. Tanaka (1983). "Economic Factors And The Composition Of Juvenile Property Crimes." *Applied Economics.* 15 (2), 267–281.

Giacopassi, D. and D. R. Forde (2000). "Broken Windows, Crumpled Fenders, And Crime." *Journal Of Criminal Justice.* 28 (5), 397–405.

Gibbs, J. P. and G. Firebaugh (1990). "The Artifact Issue In Deterrence Research." *Criminology.* 28 (2), 347–367.

Gillis, A. R. (1989). "Crime And State Surveillance In Nineteenth-Century France." *American Journal Of Sociology.* 95 (2), 307–341.

Gius, M. (1999). "The Economics Of The Criminal Behavior Of Young Adults: Estimation Of An Economic Model Of Crime With A Correction For Aggregate Market And Public Policy Variables." *American Journal Of Economics And Sociology.* 58 (4), 947–957.

Glaeser, E. L. (2006). "Researcher Incentives And Empirical Methods." Discussion Paper 2122, Harvard Institute Of Economics Research, Harvard University, Cambridge.

——— and B. Sacerdote (1999). "Why Is There More Crime In Cities?" *Journal Of Political Economy.* 107 (6), 225–258.

Glaser, D. and M. S. Zeigler (1974). "Use Of The Death Penalty Versus Outrage At Murder." *Crime And Delinquency.* 20 (4), 329–338.

Glass, G. V. (1968). "Analysis Of Data On The Connecticut Speeding Crackdown As A Time-Series Quasi-Experiment." *Law And Society Review.* 3 (1), 55–76.

Gneezy, U. and A. Rustichini (2000). "A Fine Is A Price." *Journal Of Legal Studies.* 29 (1), 1–17.

Goel, R. K. and D. P. Rich (1989). "On The Economic Incentives For Taking Bribes." *Public Choice.* 61 (3), 269–275.

——— and M. A. Nelson (1998). "Corruption And Government Size: A Disaggregated Analysis." *Public Choice.* 97 (1–2), 107–120.

Goldberg, I. and F. C. Nold (1980). "Does Reporting Deter Burglars? An Empirical Analysis Of Risk And Return In Crime." *Review Of Economics And Statistics.* 62 (3), 424–431.

Good, D. H. and M. A. Pirog-Good (1987). "Employment, Crime, And Race." *Contemporary Economic Policy.* 5 (4), 91–104.

Good, I. J. (1952). "Rational Decisions." *Journal Of The Royal Statistical Society B.* 14 (1), 107–114.

Goodman, D. E. (1997). "Midsize Cities And Their Correlates With Crime: An Empirical Investigation." Working Paper 97–01, Department Of Economics, University Of Puget Sound, Tacoma.

Gottfredson, M. R. and T. Hirschi (1987). "The Positive Tradition." in M. R. Gottfredson and T. Hirschi (eds.), *Positive Criminology*, Sage, Newbury Park, pp. 9–22.

Grasmick, H. G. and E. Kobayashi (2002). "Workplace Deviance In Japan: Applying An Extended Model Of Deterrence." *Deviant Behavior.* 23 (1), 21–44.

——— and G. J. Bryjak (1980). "The Deterrent Effect Of Perceived Severity Of Punishment." *Social Forces.* 59 (2), 471–491.

——— and H. Milligan (1976). "Deterrence Theory Approach To Socioeconomic/Demographic Correlates Of Crime." *Social Science Quarterly.* 57 (3), 608–617.

——— and L. Appleton (1977). "Legal Punishment And Social Stigma: A Comparison Of Two Deterrence Models." *Social Science Quarterly.* 58 (1), 15–28.

——— and R. J. Bursik (1990). "Conscience, Significant Others, And Rational Choice: Extending The Deterrence Model." *Law And Society Review.* 24 (3), 837–861.

——— and W. J. Scott (1982). "Tax Evasion And Mechanisms Of Social Control: A Comparison With Grand And Petty Theft." *Journal Of Economic Psychology.* 2 (3), 213–230.

———, D. Jacobs, and C. McCollom (1983). "Social Class And Social-Control: An Application Of Deterrence Theory." *Social Forces.* 62 (2), 359–374.

Grau, V. M. and T. Groves (1997). "The Oil Spill Process: The Effect Of Coast Guard Monitoring On Oil Spills." *Environmental And Resource Economics.* 10 (4), 315–339.

Gray, L. N. and D. J. Martin (1969). "Punishment And Deterrence: Another Analysis Of Gibbs' Data." *Social Science Quarterly.* 50 (2), 389–395.

Green, D. E. (1989a). "Measures Of Illegal Behavior In Individual-Level Deterrence Research." *Journal Of Research In Crime And Delinquency.* 26 (3), 253–275.

——— (1989b). "Past Behavior As A Measure Of Actual Future Behavior: An Unresolved Issue In Perceptual Deterrence Research." *Journal Of Criminal Law And Criminology.* 80 (3), 781–804.

Greenberg, D. F. (1985). “Age, Crime, And Social Explanation.” *American Journal Of Sociology.* 91 (1), 1–21.

——— and R. C. Kessler (1982a). “The Effect Of Arrests On Crime: A Multivariate Panel Analysis.” *Social Forces.* 60 (3), 771–790.

——— and R. C. Kessler (1982b). “Model Specification In Dynamic Analyses Of Crime Deterrence.” in J. Hagan (ed.), *Deterrence Reconsidered: Methodological Innovations*, Sage Publications, Beverly Hills, pp. 15–32.

———, R. C. Kessler, and C. H. Logan (1979). “A Panel Model Of Crime Rates And Arrest Rates.” *American Sociological Review.* 44 (5), 843–850.

———, R. C. Kessler, and C. Loftin (1983). “The Effect Of Police Employment On Crime.” *Criminology.* 21 (3), 375–394.

Greenwood, M. J. and W. J. Wadycki (1973). “Crime Rates And Public Expenditures For Police Protection: Their Interaction.” *Review Of Social Economy.* 31 (2), 138–151.

——— and W. J. Wadycki (1975). “Crime Rates And Public Expenditures For Police Protection: A Reply.” *Review Of Social Economy.* 33 (1), 81–85.

Grogger, J. (1990). “The Deterrent Effect Of Capital Punishment: An Analysis Of Daily Homicide Counts.” *Journal Of The American Statistical Association.* 85 (410), 295–303.

Gross, M. and S. Hakim (1982). “The Diffusion Of Crime In The Metropolis: A Step-By-Step Analysis.” *Journal Of Environmental Systems.* 12 (1), 45–60.

Grosvenor, D., T. L. Toomey, and A. C. Wagenaar (1999). “Deterrence And The Adolescent Drinking Driver.” *Journal Of Safety Research.* 30 (3), 187–191.

Gunnison, E. K. (2002). “Understanding Female Desistance From Crime: Exploring Theoretical And Empirical Relationships.” PhD dissertation, University of Cincinnati, Cincinnati.

Guttel, E. and B. Medina (2007). “Less Crime, More (Vulnerable) Victims: Game Theory And The Distributional Effects Of Criminal Sanctions.” *Review Of Law And Economics.* 3 (2), 407–435.

Gyimah-Brempong, K. (1986). “Empirical Models Of Criminal Behavior: How Significant A Factor Is Race?” *Review Of Black Political Economy.* 15 (1), 27–43.

Hakim, S. (1980). “The Attraction Of Property Crimes To Suburban Localities: A Revised Economic Model.” *Urban Studies.* 17 (3), 265–276.

———, A. Ovadia, and J. Weinblatt (1978). “Crime Attraction And Deterrence In Small Communities: Theory And Results.” *International Regional Science Review.* 3 (2), 153–163.

———, A. Ovadia, E. Sagi, and J. Weinblatt (1979). “Interjurisdictional Spillover Of Crime And Police Expenditure.” *Land Economics.* 55 (2), 200–212.

Hale, C. (1998). “Crime And The Business Cycle In Post-War Britain Revisited.” *British Journal Of Criminology.* 38 (3), 681–698.

——— (1999). “The Labour Market And Post-War Crime Trends In England And Wales.” in P. Carlen and R. Morgan (eds.), *Crime Unlimited? Questions For The 21st Century*, MacMillan, Basingstoke.

Hannan, T. H. (1982). “Bank Robberies And Bank Security Precautions.” *Journal Of Legal Studies.* 11 (1), 83–92.

Hansen, B. E. (2007). “Least Squares Model Averaging.” *Econometrica.* 75 (4), 1175–1189.

Hansen, K. and S. Machin (2002). “Spatial Crime Patterns And The Introduction Of The UK Minimum Wage.” *Oxford Bulletin Of Economics And Statistics.* 64 (Suppl. S), 677–697.

Haque, O. and M. Cameron (1989). “Effect Of The Victorian Zero BAC Legislation On Serious Casualty Accidents: July 1984 - December 1995.” *Journal Of Safety Research.* 20 (3), 129–137.

Harcourt, B. E. and J. Ludwig (2006). “Broken Windows: New Evidence From New York City And A Five-City Social Experiment.” *University Of Chicago Law Review.* 73 (1), 271–320.

Hashimoto, M. (1987). “The Minimum-Wage Law And Youth Crimes - Time-Series Evidence.” *Journal Of Law And Economics.* 30 (2), 443–464.

Hatcher, A., S. Jaffry, O. Thébaud, and E. Bennett (1999). “Factors Affecting Compliance With Fishery Regulations: A UK Case-Study.” in P. Salz (ed.), *Proceedings Of The Xth Annual Conference Of The European Association Of Fisheries Economists*, Agricultural Economics Research Institute, The Hague, pp. 157–172.

Heckelman, J. C. and A. C. Yates (2003). “And A Hockey Game Broke Out: Crime And Punishment In The NHL.” *Economic Inquiry.* 41 (4), 705–712.

Hedges, L. V. and I. Olkin (1985). *Statistical Methods For Meta-Analysis*, Academic Press, San Diego.

Heineke, J. M. (1978a). *Economic Models Of Criminal Behaviour*, North Holland Publishing Company, Amsterdam.

——— (1978b). “Substitution Among Crimes And The Question Of Deterrence: An Indirect Utility Function Approach To The Supply Of Legal And Illegal Activity.” in J. M. Heineke (ed.), *Economic Models Of Criminal Behaviour*, North Holland Publishing Company, Amsterdam, pp. 153–210.

Heisler, G. H. (1974). “Ways To Deter Law Violators: Effects Of Levels Of Threat And Vicarious Punishment On Cheating.” *Journal Of Consulting And Clinical Psychology.* 42 (4), 577–582.

Hendrey, D. F. and H. M. Krolzig (2000). “Computer Automation Of General-To-Specific Model Selection Procedures.” Working Paper, Institute Of Economics And Statistics, Nuffield College, Oxford.

Hessing, D. J., H. Elffers, and R. H. Weigel (1988). “Exploring The Limits Of Self-Reports And Reasoned Action: An Investigation Of The Psychology Of Tax Evasion Behavior.” *Journal Of Personality And Social Psychology.* 54 (3), 405–413.

———, H. Elffers, H. S. Robben, and P. Webley (1992). “Does Deterrence Deter? Measuring The Effect Of Deterrence On Tax Compliance In Field Studies And Experimental Studies.” in J. Slemrod (ed.), *Why People Pay Taxes: Tax Compliance And Enforcement*, University Of Michigan Press, Ann Arbor, pp. 291–310.

Hill, J. P. and R. A. Kochendorfer (1969). "Knowledge Of Peer Success And Risk Of Detection As Determinants Of Cheating." *Developmental Psychology.* 1 (3), 231–238.

Hilton, M. E. (1984). "The Impact Of Recent Changes In California Drinking-Driving Laws On Fatal Accident Levels During The First Postintervention Year: An Interrupted Time Series Analysis." *Law And Society Review.* 18 (4), 605–627.

Hingson, R., T. Heeren, and M. Winter (1996). "Lowering State Legal Blood Alcohol Limits To 0.08%: The Effect On Fatal Motor Vehicle Crashes." *American Journal Of Public Health.* 86 (9), 1297–1299.

Hirschi, T. and M. Gottfredson (1983). "Age And The Explanation Of Crime." *American Journal Of Sociology.* 89 (3), 552–584.

——— and M. Gottfredson (1985). "Age And Crime, Logic And Scholarship: Comment On Greenberg." *American Journal Of Sociology.* 91 (1), 22–27.

Hite, P. A. (1988). "The Effect Of Peer Reporting Behavior On Taxpayer Compliance." *Journal Of The American Taxation Association.* 9 (2), 47–64.

Hoenack, S. A. and W. C. Weiler (1980). "A Structural Model Of Murder Behavior And The Criminal Justice System." *American Economic Review.* 70 (3), 327–341.

——— , R. T. Kudrle, and D. L. Sjoquist (1978). "The Deterrent Effect Of Capital Punishment: A Question Of Identification." *Policy Analysis.* 4 (4), 491–527.

Hoeting, J. A., D. M. Adrian, E. Raftery, and C. T. Volinsky (1999). "Bayesian Model Averaging: A Tutorial (With Discussion)." *Statistical Science.* 14 (4), 382–417.

Hollinger, R. C. and J. P. Clark (1983). "Deterrence In The Workplace: Perceived Certainty, Perceived Severity And Employee Theft." *Social Forces.* 62 (2), 398–418.

Holtmann, A. G. and L. Yap (1978). "Does Punishment Pay?" *Public Finance.* 33 (1–2), 90–97.

Homel, R. J. (1994). "Drink-Driving Law Enforcement And The Legal Blood Alcohol Limit In New South Wales." *Accident Analysis And Prevention.* 26 (2), 147–155.

Hoover, K. D. and S. J. Perez (1999). "Data Mining Reconsidered: Encompassing And The General-To-Specific Approach To Specification Search." *Econometrics Journal.* 2 (2), 167–191.

Horai, J. and J. T. Tedeschi (1969). "Effects Of Credibility And Magnitude Of Punishment On Compliance To Threats." *Journal Of Personality And Social Psychology.* 12 (2), 164–169.

Houston, D. J. and L. E. Richardson (2004). "Drinking And Driving In America: A Test Of Behavioral Assumptions Underlying Public Policy." *Political Research Quarterly.* 57 (1), 53–64.

Houston, J. P. (1983). "College Classroom Cheating, Threat, Sex And Prior Performance." *College Student Journal.* 17 (3), 229–235.

Howsen, R. M. and S. B. Jarrell (1987). "Some Determinants Of Property Crime: Economic Factors Influence Criminal Behavior But Cannot Completely Explain The Syndrome." *American Journal Of Economics And Sociology.* 46 (4), 445–457.

Hsing, Y. (1996). “An Analysis Of Arrests Regarding Illegal Drugs: The Determinants And Policy Implications.” *American Journal Of Economics And Sociology.* 55 (1), 53–60.

Huff, R. C. and J. M. Stahura (1980). “Police Employment And Suburban Crime.” *Criminology.* 17 (4), 461–470.

Hull, B. B. and F. Bold (1995). “Preaching Matters: Replication And Extension.” *Journal Of Economic Behavior And Organization.* 27 (1), 143–149.

Humphries, D. and D. Wallace (1980). “Capitalist Accumulation And Urban Crime.” *Social Problems.* 28 (2), 179–193.

Hurst, P. M. (1978). “Blood Test Legislation In New Zealand.” *Accident Analysis And Prevention.* 10 (4), 287–296.

Ihlanfeldt, K. R. (2003). “Rail Transit And Neighborhood Crime: The Case Of Atlanta, Georgia.” *Southern Economic Journal.* 70 (2), 273–294.

Isachsen, A. J., S. O. Samuelson, and S. Strøm (1985). “The Behavior Of Tax Evaders.” in A. Wenig and W. Gaertner (eds.), *The Economics Of The Shadow Economy*, Springer, Berlin, pp. 227–244.

Jackson, B. R. and S. M. Jones (1985). “Salience Of Tax Evasion Penalties Versus Detection Risk.” *Journal Of The American Taxation Association.* 6 (2), 7–17.

Jacob, H. and M. J. Rich (1981). “The Effects Of The Police On Crime: A Second Look.” *Law And Society Review.* 15 (1), 109–122.

Jarrell, S. B. and R. M. Howsen (1990). “Transient Crowding And Crime: The More Strangers In An Area, The More Crime Except For Murder, Assault And Rape.” *American Journal Of Economics And Sociology.* 49 (4), 483–494.

Jayewardene, C. H. (1973). “Life Or Death: Society’s Reaction To Murder.” *Canadian Journal Of Criminology And Corrections.* 15 (3), 265–273.

Jensen, E. K. and L. K. Metsger (1994). “A Test Of The Deterrent Effect Of Legislative Waiver On Violent Juvenile Crime.” *Crime And Delinquency.* 40 (1), 96–104.

Jensen, G. F. (1969). “Crime Doesn’t Pay: Correlates Of A Shared Misunderstanding.” *Social Problems.* 17 (2), 189–201.

——— and B. G. Stitt (1982). “Words And Misdeeds: Hypothetical Choice Versus Past Behavior As Measures Of Deviance.” in J. Hagan (ed.), *Deterrence Reconsidered*, Beverly Hills, pp. 33–54.

——— , M. L. Erickson, and J. P. Gibbs (1978). “Perceived Risk Of Punishment And Self-Reported Delinquency.” *Social Forces.* 57 (1), 57–78.

Johnson, B. T., B. Mullen, and E. Salas (1995). “Comparison Of Three Major Meta-Analytic Approaches.” *Journal Of Applied Psychology.* 80 (1), 94–106.

Johnson, D. and J. C. Fell (1995). “The Impact Of Lowering The Illegal BAC Limit To .08 In Five States In The U.S.” in “39th Annual Proceedings Of The Association For The Advancement Of Automotive Medicine,” Chicago.

Johnson, J. L. and K. Lesniak-Karpiak (1997). "The Effect Of Warning On Malingering On Memory And Motor Tasks In College Samples." *Archives of Clinical Neuropsychology.* 12 (3), 231–238.

Jones, T. E. (1973). "Evaluating Everyday Policies: Police Activity And Crime Incidence." *Urban Affairs Quarterly.* 8 (3), 267–279.

Kahan, D. M. (2004). "The Theory Of Value Dilemma: A Critique Of The Economic Analysis Of Criminal Law." *Ohio State Journal Of Criminal Law.* 1 (2), 643–651.

Kalfus, G. R., J. R. Ferrari, P. Arean, and D. Balser (1987). "An Examination Of The New York Mandatory Seat Belt Law On A University Campus." *Law And Human Behavior.* 11 (1), 63–67.

Kaminski, R. J. and T. B. Marvell (2002). "A Comparison Of Changes In Police And General Homicides: 1930-1998." *Criminology.* 40 (1), 171–190.

Kaplan, H. B. and K. R. Damphousse (1997). "Negative Social Sanctions, Self Derogation, And Deviant Behavior: Main And Interactive Effects In Longitudinal Perspective." *Deviant Behavior.* 18 (1), 1–26.

Kaplan, J. (1983). *The Hardest Drug: Heroin And Public Policy*, University Of Chicago Press, Chicago.

Karpoff, J. M., J. R. Lott, and G. Rankine (1998). "Environmental Violations, Legal Penalties, And Reputation Costs." Working Paper 71, University Of Chicago Law School, Chicago.

Karstedt-Henke, S. (1991). "Attribution Theory And Deterrence Research: A New Approach To Old Problems." in K. Sessar and H.-J. Kerner (eds.), *Developments In Crime And Crime Control Research. German Studies On Victims, Offenders And The Public*, Springer, Berlin, pp. 22–40.

Katz, L., S. D. Levitt, and E. Shustorovich (2003). "Prison Conditions, Capital Punishment, And Deterrence." *American Law And Economics Review.* 5 (2), 318–343.

Katzman, M. T. (1980). "The Contribution Of Crime To Urban Decline." *Urban Studies.* 17 (3), 277–286.

Kau, J. B. and P. H. Rubin (1975). "New Estimates Of The Determinants Of Urban Crime." *Annals Of Regional Science.* 9 (1), 68–76.

Keane, C., A. R. Gillis, and J. Hagan (1989). "Deterrence And Amplification Of Juvenile Delinquency By Police Contact: The Importance Of Gender And Risk-Orientation." *British Journal Of Criminology.* 29 (4), 336–352.

Kelling, G. L. and W. H. Sousa (2001). "Do Police Matter? An Analysis Of The Impact Of New York City's Police Reforms." Civic Report 22, Center For Civic Innovation, The Manhattan Institute For Policy Research, New York.

Kelly, M. (2000). "Inequality And Crime." *Review Of Economics And Statistics.* 82 (4), 530–539.

Kenkel, D. S. (1993). "Drinking, Driving, And Deterrence: The Effectiveness And Social Costs Of Alternative Policies." *Journal Of Law And Economics.* 36 (2), 877–913.

——— and S. F. Koch (2001). "Deterrence And Knowledge Of The Law: The Case Of Drunk Driving." *Applied Economics.* 33 (7), 845–854.

Kessler, D. P. and S. D. Levitt (1998). "Using Sentence Enhancements To Distinguish Between Deterrence And Incapacitation." *Journal Of Law And Economics.* 42 (1), 343–363.

Kessler, M. and A. Molinari (1997). "Una Aproximación Microeconómica Al Crimen En La Argentina." Working Paper 1248, Asociacion Argentina De Economia Politica, Buenos Aires.

Killias, M. (1985). "Zur Bedeutung von Rechtsgefühl und Sanktionen für die Konformität des Verhaltens gegenüber neuen Normen: Das Beispiel der Gurtanlegepflicht." in E. J. Lampe (ed.), *Jahrbuch für Rechtssoziologie und Rechtstheorie*, Vol. 10, Opladen.

Kim, I., B. L. Benson, D. W. Rasmussen, and T. W. Zuehlke (1993). "An Economic Analysis Of Recidivism Among Drug Offenders." *Southern Economic Journal.* 60 (1), 169–183.

Kim, K. S., M. H. Myeong, and K. Young-Jun (2006). "Evaluating The Effects Of Safety Policy Measures On Traffic Fatalities In Korea." *Transport Reviews.* 26 (3), 293–304.

Kirchner, R. E., J. F. Schnelle, M. Domash, L. Larson, A. Carr, and M. P. McNees (1980). "The Applicability Of A Helicopter Patrol Procedure To Diverse Areas: A Cost-Benefit Evaluation." *Journal Of Applied Behavior Analysis.* 13 (1), 143–148.

Kleck, G. (1979). "Capital Punishment, Gun Ownership, And Homicide." *American Journal Of Sociology.* 84 (4), 882–910.

Klemke, L. W. (1978). "Does Apprehension For Shoplifting Amplify Or Terminate Shoplifting?" *Law And Society Review.* 12 (3), 391–403.

Klepper, S. and D. Nagin (1989a). "The Deterrent Effect Of Perceived Certainty And Severity Of Punishment Revisited." *Criminology.* 27 (4), 721–746.

——— and D. Nagin (1989b). "Tax Compliance And Perceptions Of The Risk Of Detection And Criminal Prosecution." *Law And Society Review.* 23 (2), 209–240.

——— and D. S. Nagin (1989c). "The Anatomy Of Tax Evasion." *Journal Of Law, Economics And Organization.* 5 (1), 1–24.

Klick, J. and A. Tabarrok (2005). "Using Terror Alert Levels To Estimate The Effect Of Police On Crime." *Journal Of Law And Economics.* 48 (1), 267–279.

Knell, M. and H. Stix (2005). "The Income Elasticity Of Money Demand: A Meta-Analysis Of Empirical Results." *Journal Of Economic Surveys.* 19 (3), 513–533.

Kohfeld, C. W. (1983). "Rational Cops, Rational Robbers, And Information." *Journal Of Criminal Justice.* 11 (5), 459–466.

Konar, S. and M. A. Cohen (2000). "Why Do Firms Pollute (And Reduce) Toxic Emissions?" Working Paper, Owen School of Management, Vanderbilt University, Nashville, Tennessee.

Koop, G. and S. Potter (2003). "Forecasting In Large Macroeconomic Panels Using Bayesian Model Averaging." Staff Report 163, Federal Reserve Bank, New York.

Koskela, E. and M. Virén (1993). "An Economic Model Of Auto Thefts In Finland." *International Review Of Law And Economics.* 13 (2), 179–191.

——— and M. Virén (1994). "Auto Thefts In Finland: An Occupational Choice Model." *Public Finance.* 49 (Suppl. S), 168–181.

——— and M. Virén (1997). "An Occupational Choice Model Of Crime Switching." *Applied Economics.* 29 (5), 655–660.

Kovandzic, T. V. and J. J. Sloan (2002). "Police Levels And Crime Rates Revisited: A County-Level Analysis From Florida (1980-1998)." *Journal Of Criminal Justice.* 30 (1), 65–76.

——— and T. B. Marvell (2003). "Right-To-Carry Concealed Handguns And Violent Crime: Crime Control Through Gun Decontrol?" *Criminology And Public Policy.* 2 (3), 363–396.

———, J. J. Sloan, and L. M. Vieraitis (2004). "Striking Out As Crime Reduction Policy: The Impact Of Three Strikes Laws On Crime Rates In U.S. Cities." *Justice Quarterly.* 21 (2), 207–239.

Krakovsky, M. (2004). "Register Or Perish." *Scientific American.* 291 (6), 18–20.

Krylo, D. J. (1985). "The Effect Of Uniformed Police Officers On The Deterrence Of Crime." PhD dissertation.

Kuperan, K. and J. G. Sutinen (1998). "Blue Water Crime: Deterrence, Legitimacy, And Compliance In Fisheries." *Law And Society Review.* 32 (2), 309–337.

Land, K. C. and M. Felson (1976). "A General Framework For Building Dynamic Macro Social Indicator Models: Including An Analysis Of Changes In Crime Rates And Police Expenditures." *American Journal Of Sociology.* 82 (3), 565–604.

Landes, W. M. (1978). "An Economic Study Of U.S. Aircraft Hijacking, 1961–1976." *Journal Of Law And Economics.* 21 (1), 1–31.

Laycock, G. (1991). "Operation Identification, Or The Power Of Publicity?" *Security Journal.* 2 (2), 67–72.

Layson, S. K. (1983). "Homicide And Deterrence: Another View Of The Canadian Time-Series Evidence." *Canadian Journal Of Economics-Revue Canadienne De Economique.* 16 (1), 52–73.

——— (1985). "Homicide And Deterrence: A Reexamination Of The United States Time-Series Evidence." *Southern Economic Journal.* 52 (1), 68–89.

Layton, A. P. (1983). "The Impact Of Increased Penalties On Australian Drink/Driving Behaviour." *Logistics And Transportation Review.* 19 (3), 261–266.

Leamer, E. E. (1982). "Sets Of Posterior Means With Bounded Variance Priors." *Econometrica.* 50 (3), 725–736.

——— (1983). "Let's Take The Con Out Of Econometrics." *American Economic Review.* 73 (1), 31–43.

——— (1985). "Sensitivity Analyses Would Help." *American Economic Review.* 75 (3), 308–313.

Lee, D. and J. McCrary (2005). "Crime, Punishment, And Myopia." Working Paper No. W11491, National Bureau of Economic Research, Cambridge, Massachusetts.

Legge, J. S. (1990). "Reforming Highway Safety In New York State: An Evaluation Of Alternative Policy Interventions." *Social Science Quarterly.* 71 (2), 373–382.

——— and J. Park (1994). "Policies To Reduce Alcohol-Impaired Driving: Evaluating Elements Of Deterrence." *Social Science Quarterly.* 75 (3), 594–606.

Lenton, S. (2000). "Cannabis Policy And The Burden Of Proof: Is It Now Beyond Reasonable Doubt That Cannabis Prohibition Is Not Working?" *Drug And Alcohol Review.* 19 (1), 95–100.

Leung, A. (2002). "Delinquency, Social Institutions, And Capital Accumulation." *Journal Of Institutional And Theoretical Economics - Zeitschrift für die gesamte Staatswissenschaft.* 158 (3), 420–440.

Levine, R. and D. Renelt (1992). "A Sensitivity Analysis Of Cross-Country Growth Regressions." *American Economic Review.* 82 (4), 942–963.

Levitt, S. D. (1996). "The Effect Of Prison Population Size On Crime Rates: Evidence From Prison Overcrowding Litigation." *Quarterly Journal Of Economics.* 111 (2), 319–351.

——— (1997a). "Incentive Compatibility Constraints As An Explanation For The Use Of Prison Sentences Instead Of Fines." *International Review Of Law And Economics.* 17 (2), 179–192.

——— (1997b). "Using Electoral Cycles In Police Hiring To Estimate The Effect Of Police On Crime." *American Economic Review.* 87 (3), 270–290.

——— (1998a). "Juvenile Crime And Punishment." *Journal Of Political Economy.* 106 (6), 1156–1185.

——— (1998b). "The Relationship Between Crime Reporting And Police: Implications For The Use Of Uniform Crime Reports." *Journal Of Quantitative Criminology.* 14 (1), 61–81.

——— (1998c). "Why Do Increased Arrest Rates Appear To Reduce Crime: Deterrence, Incapacitation, Or Measurement Error?" *Economic Inquiry.* 36 (3), 353–372.

——— (2001). "Alternative Strategies For Identifying The Link Between Unemployment And Crime." *Journal Of Quantitative Criminology.* 17 (4), 377–390.

——— (2002a). "Testing The Economic Model Of Crime: The National Hockey League's Two-Referee Experiment." *Contributions To Economic Analysis And Policy.* 1 (1). Article 2.

——— (2002b). "Using Electoral Cycles In Police Hiring To Estimate The Effects Of Police On Crime: Reply." *American Economic Review.* 92 (4), 1244–1250.

——— (2004). "Understanding Why Crime Fell In The 1990s: Four Factors That Explain The Decline And Six That Do Not." *Journal Of Economic Perspectives.* 18 (1), 163–190.

——— and J. Porter (2001). "How Dangerous Are Drinking Drivers?" *Journal Of Political Economy.* 109 (6), 1198–1237.

——— and L. Lochner (2001). "The Determinants Of Juvenile Crime." in J. Gruber (ed.), *Risky Behavior Among Youths: An Economic Analysis*, The University Of Chicago Press, Chicago, pp. 327–373.

Levy, P., R. B. Voas, P. Johnson, and T. Klein (1978). "An Evaluation Of The Department Of Transportation's Alcohol Safety Action Projects." *Journal Of Safety Research.* 10 (4), 162–176.

Liang, F., Y. K. Truong, and W. H. Wong (2001). "Automatic Bayesian Model Averaging For Linear Regression And Applications In Bayesian Curve Fitting." *Statistica Sinica.* 11 (4), 1005–1029.

Liu, Y. W. and R. H. Bee (1983). "Modeling Criminal Activity In An Area In Economic Decline: Local Economic Conditions Are A Major Factor In Local Property Crime." *American Journal Of Economics And Sociology.* 42 (2), 385–392.

Lochner, L. (2001). "A Theoretical And Empirical Study Of Individual Perceptions Of The Criminal Justice System." Working Paper 483, Center For Economic Research, University Of Rochester, Rochester.

——— (2003). "Individual Perceptions Of The Criminal Justice System." Working Paper 9474, National Bureau of Economic Research, Cambridge, Massachusetts.

——— (2004). "Education, Work, And Crime: A Human Capital Approach." *International Economic Review.* 45 (3), 811–843.

Loftin, C. (1980). "The Deterrent Effect Of Punishment." in S. E. Fienberg and A. J. Reiss (eds.), *Indicators Of Crime And Criminal Justice: Quantitative Studies*, U.S. Department Of Justice, Washington D.C., pp. 75–81.

——— and D. McDowall (1982). "The Police, Crime, And Economic Theory: An Assessment." *American Sociological Review.* 47 (3), 393–401.

Logan, C. H. (1972). "General Deterrent Effects Of Imprisonment." *Social Forces.* 51 (1), 64–73.

——— (1973). "Legal Sanctions And Deterrence From Crime." PhD dissertation, Indiana University, Bloomington.

——— (1975). "Arrest Rates And Deterrence." *Social Science Quarterly.* 56 (3), 376–389.

Longhi, S., P. Nijkamp, and J. Poot (2005). "A Meta-Analytic Assessment Of The Effect Of Immigration On Wages." *Journal Of Economic Surveys.* 19 (3), 451–477.

Lott, J. R. (1987). "Juvenile Delinquency And Education: A Comparison Of Public And Private Provision." *International Review Of Law And Economics.* 7 (2), 163–175.

——— (1998). "The Concealed-Handgun Debate." *Journal Of Legal Studies.* 27 (1), 221–243.

——— and D. B. Mustard (1997). "Crime, Deterrence, And Right-To-Carry Concealed Handguns." *Journal Of Legal Studies.* 26 (1), 1–68.

——— and J. E. Whitley (2001). "Safe-Storage Gun Laws: Accidental Deaths, Suicides, And Crime." *Journal Of Law And Economics.* 44 (2), 659–689.

Lovell, M. C. (1983). "Data Mining." *Review Of Economics And Statistics.* 65 (1), 1–12.

Low, S. A. and L. R. McPheters (1983). "Wage Differentials And Risk Of Death: An Empirical Analysis." *Economic Inquiry.* 21 (2), 271–280.

Lu, H., L. Zhang, and T. D. Miethe (2002). "Interdependency, Communitarianism And Reintegrative Shaming In China." *Social Science Journals.* 39 (2), 189–201.

Ludwig, J. (1998). "Concealed-Gun-Carrying Laws And Violent Crime: Evidence From State Panel Data." *International Review Of Law And Economics.* 18 (3), 239–254.

Luiz, J. M. (2001). "Temporal Association, The Dynamics Of Crime, And Their Economic Determinants: A Time Series Econometric Model Of South Africa." *Social Indicators Research.* 53 (1), 33–61.

Lynch, M. J. (1999). “Beating A Dead Horse: Is There Any Basic Empirical Evidence For The Deterrent Effect Of Imprisonment?” *Crime Law And Social Change.* 31 (4), 347–362.

MacDonald, J. M. (1999). “Violence And Drug Use In Juvenile Institutions.” *Journal Of Criminal Justice.* 27 (1), 33–44.

Machin, S. and C. Meghir (2004). “Crime And Economic Incentives.” *Journal Of Human Resources.* 39 (4), 958–979.

Magat, W. A. and K. W. Viscusi (1990). “Effectiveness Of The EPA’s Regulatory Enforcement: The Case Of Industrial Effluent Standards.” *Journal Of Law And Economics.* 33 (2), 331–360.

Maghsoodloo, S., D. B. Brown, and P. A. Greathouse (1988). “Impact Of The Revision Of DUI Legislation In Alabama.” *American Journal Of Drug And Alcohol Abuse.* 14 (1), 97–108.

Mann, R. E., E. R. Vingilis, D. Gavin, E. Adlaf, and L. Anglin (1991). “Sentence Severity And The Drinking Driver: Relationships With Traffic Safety Outcome.” *Accident Analysis And Prevention.* 23 (6), 483–491.

——— , R. G. Smart, and L. Anglin (1996). “Alcohol-Related Measures As Factors In Traffic Fatalities.” *Journal Of Studies On Alcohol.* 57 (6), 646–651.

——— , R. G. Smart, G. Stoduto, D. Beirness, R. Lamble, and E. Vingilis (2002). “The Early Effects Of Ontario’s Administrative Driver’s Licence Suspension Law On Driver Fatalities With A BAC $>$ 80 mg%.” *Revue Canadienne Santé Publique.* 93 (3), 176–180.

——— , R. G. Smart, G. Stoduto, E. M. Adlaf, E. Vingilis, D. Beirness, R. Lamble, and M. Asbridge (2003). “The Effects Of Drinking-Driving Laws: A Test Of The Differential Deterrence Hypothesis.” *Addiction.* 98 (11), 1531–1536.

——— , S. McDonald, G. Stoduto, S. Bondy, B. Jonah, and A. Shaikh (2001). “The Effects Of Introducing Or Lowering Legal Per Se Blood Alcohol Limits For Driving: An International Review.” *Accident Analysis And Prevention.* 33 (5), 569–583.

Marchese Beccaria, C. B. (1819). *Of Crimes And Punishment*, Heward, Robert, London. Published 1764 In Italian.

Marselli, R. and M. Vannini (1997). “Estimating A Crime Equation In The Presence Of Organized Crime: Evidence From Italy.” *International Review Of Law And Economics.* 17 (1), 89–113.

Marvell, T. B. and C. E. Moody (1991). “Age Structure And Crime Rates: The Conflicting Evidence.” *Journal Of Quantitative Criminology.* 7 (3), 237–273.

——— and C. E. Moody (1994). “Prison Population Growth And Crime Reduction.” *Journal Of Quantitative Criminology.* 10 (2), 109–140.

——— and C. E. Moody (1996). “Specification Problems, Police Levels And Crime Rates.” *Criminology.* 34 (4), 609–646.

——— and C. E. Moody (1997). “The Impact Of Prison Growth On Homicide.” *Homicide Studies.* 1 (3), 205–233.

——— and C. E. Moody (2001). “The Lethal Effects Of Three-Strikes Laws.” *Journal Of Legal Studies.* 30 (1), 89–106.

Masih, A. M. and R. Masih (1996). "Temporal Causality And The Dynamics Of Different Categories Of Crime And Their Socioeconomic Determinants: Evidence From Australia." *Applied Economics.* 28 (9), 1093–1104.

Masih, R. (1995). "Modelling The Dynamic Interactions Among Crime, Deterrence And Socio-Economic Variables: Evidence From A Vector Error-Correction Model." *Mathematics And Computers In Simulation.* 39 (3–4), 411–416.

Mason, R. and L. D. Calvin (1978). "A Study Of Admitted Income Tax Evasion." *Law And Society Review.* 13 (1), 73–89.

Mast, B. D., B. L. Benson, and D. W. Rasmussen (2000). "Entrepreneurial Police And Drug Enforcement Policy." *Public Choice.* 104 (3–4), 285–308.

Mathieson, D. and P. Passell (1976). "Homicide And Robbery In New York City: An Economic Model." *Journal Of Legal Studies.* 5 (1), 83–98.

Mathur, V. K. (1978). "Economics Of Crime: An Investigation Of The Deterrent Hypothesis For Urban Areas." *Review Of Economics And Statistics.* 60 (3), 459–466.

Matsueda, R. L., D. A. Kreager, and D. Huizinga (2006). "Deterring Delinquents: A Rational Choice Model Of Theft And Violence." *American Sociological Review.* 71 (1), 95–122.

Mauser, G. A. and D. Maki (2003). "An Evaluation Of The 1977 Canadian Firearm Legislation: Robbery Involving A Firearm." *Applied Economics.* 35 (4), 423–436.

May, P. J. (2005). "Regulation And Compliance Motivations: Examining Different Approaches." *Public Administration Review.* 65 (1), 31–44.

McAleer, M. and M. R. Veall (1989). "How Fragile Are Fragile Inferences? A Re-Evaluation Of The Deterrent Effect Of Capital Punishment." *Review Of Economics And Statistics.* 71 (1), 99–106.

McCarthy, B. and J. Hagan (2005). "Danger And The Decision To Offend." *Social Forces.* 83 (3), 1065–1096.

McCarthy, P. S. and W. Oesterle (1987). "The Deterrent Effects Of Stiffer DUI Laws: An Empirical Study." *Logistics Transportation Review.* 23 (4), 353–371.

McCormick, R. E. and R. D. Tollison (1984). "Crime On The Court." *Journal Of Political Economy.* 92 (2), 223–235.

McCrary, J. (2002). "Using Electoral Cycles In Police Hiring To Estimate The Effect Of Police On Crime: Comment." *American Economic Review.* 92 (4), 1236–1243.

McDowall, D., C. Loftin, and B. Wiersema (1992). "A Comparative Study Of The Preventive Effects Of Mandatory Sentencing Laws For Gun Crimes." *Journal Of Criminal Law And Criminology.* 83 (2), 378–394.

——— , C. Loftin, and B. Wiersema (1995). "Easing Concealed Firearms Laws: Effects On Homicide In Three States." *Journal Of Criminal Law And Criminology.* 86 (1), 193–206.

McFarland, S. G. (1983). "Is Capital Punishment A Short-Term Deterrent To Homicide? A Study Of The Effects Of Four Recent American Executions." *Journal Of Criminal Law And Criminology.* 74 (3), 1014–1032.

McGarrell, E. F., S. Chermak, A. Weiss, and J. Wilson (2001). “Reducing Firearms Violence Through Directed Police Patrol.” *Criminology And Public Policy.* 1 (1), 119–148.

McGeorge, J. and C. K. Aitken (1997). “Effects Of Cannabis Decriminalization In The Australian Capital Territory On University Students’ Patterns Of Use.” *Journal Of Drug Issues.* 27 (4), 785–793.

McManus, W. S. (1985). “Estimates Of The Deterrent Effect Of Capital Punishment: The Importance Of The Researcher’s Prior Beliefs.” *Journal Of Political Economy.* 93 (2), 417–425.

McPheters, L. R. (1978). “Econometric Analysis Of Factors Influencing Crime On The Campus.” *Journal Of Criminal Justice.* 6 (1), 47–52.

——— and W. B. Stronge (1974). “Law Enforcement Expenditures And Urban Crime.” *National Tax Journal.* 27 (4), 633–644.

Medoff, M. H. and J. P. Magaddino (1982). “An Empirical Analysis Of No-Fault Insurance.” *Evaluation Review.* 6 (3), 373–392.

Meera, A. K. and M. D. Jayakumar (1995). “Determinants Of Crime In A Developing Country: A Regression Model.” *Applied Economics.* 27 (5), 455–460.

Mehay, S. L. (1977). “Interjurisdictional Spillovers Of Urban Police Services.” *Southern Economic Journal.* 43 (3), 1352–1359.

——— and R. L. Pacula (1999). “The Effectiveness Of Workplace Drug Prevention Policies: Does ’zero Tolerance’ Work?” Working Paper 7383, National Bureau of Economic Research, Cambridge, Massachusetts.

Meier, K. J. (1992). “The Politics Of Drug-Abuse: Laws, Implementation, And Consequences.” *Western Political Quarterly.* 45 (1), 41–69.

Meier, R. F. (1982). “Jurisdictional Differences In Deterring Marijuana Use.” *Journal Of Drug Issues.* 12 (1), 61–71.

——— and W. T. Johnson (1977). “Deterrence As Social-Control: Legal And Extralegal Production Of Conformity.” *American Sociological Review.* 42 (2), 292–304.

———, S. R. Burkett, and C. A. Hickman (1984). “Sanctions, Peers, And Deviance: Preliminary Models Of A Social Control Process.” *Sociological Quarterly.* 25 (1), 67–82.

Mendes, S. M. (2004). “Certainty, Severity, And Their Relative Deterrent Effects: Questioning The Implications Of The Role Of Risk In Criminal Deterrence Policy.” *Policy Studies Journal.* 32 (1), 59–74.

——— and M. D. McDonald (2001). “Putting Severity Of Punishment Back In The Deterrence Package.” *Policy Studies Journal.* 29 (4), 588–610.

Merrifield, J. (1997). “Examining Substitution Between Property Crimes Using North Carolina Data.” Research Paper 9723, Regional Research Institute, National Science Foundation Undergraduate Research Fellow, West Virginia University, Morgantown.

Merriman, D. (1988). “Homicide And Deterrence: The Japanese Case.” *International Journal Of Offender Therapy And Comparative Criminology.* 32 (1), 1–16.

Merton, R. K. (1938). “Social Structure And Anomie.” *American Sociological Review.* 3 (5), 672–682.

Messner, S. F. and R. Rosenfeld (1997). *Crime And The American Dream*, Wadsworth, Balmont, California.

Meyer, S. (2007). “Sozialkapital und Delinquenz - Eine empirische Untersuchung für Deutschland.” PhD dissertation, Darmstadt University Of Technology, Darmstadt.

Michaels, R. M. (1960). “The Effects Of Enforcement On Traffic Behaviour.” *Public Roads.* 31 (5), 109–113.

Mikesell, J. and M. A. Pirog-Good (1990). “State Lotteries And Crime: The Regressive Revenue Producer Is Linked With A Crime Rate Higher By 3 Percent.” *American Journal Of Economics And Sociology.* 49 (1), 7–19.

Miller, J. L. and A. B. Anderson (1986). “Updating The Deterrence Doctrine.” *Journal Of Criminal Law And Criminology.* 77 (2), 418–438.

Minor, W. W. (1975). “Control Theory And Deterrence Of Crime: A Theoretical And Empirical Integration.” PhD dissertation, Florida State University, Tallahassee.

——— and J. Harry (1982). “Deterrent And Experiential Effects In Perceptual Deterrence Research: A Replication And Extension.” *Journal Of Research In Crime And Delinquency.* 19 (2), 190–203.

Miranne, A. C. (1981). “General And Specific Deterrence: An Experimental Assessment.” PhD dissertation, Washington State University, Pullman.

——— and L. N. Gray (1987). “Deterrence: A Laboratory Experiment.” *Deviant Behavior.* 8 (2), 191–203.

Miron, J. A. (2001). “Violence, Guns, And Drugs: A Cross-Country Analysis.” *Journal Of Law And Economics.* 44 (2), 615–633.

Mixon, F. G. and D. C. Mixon (1996). “The Economics Of Illegitimate Activities: Further Evidence.” *Journal Of Socio-Economics.* 25 (3), 373–381.

Mocan, N. H. and D. I. Rees (2005). “Economic Conditions, Deterrence And Juvenile Crime: Evidence From Micro Data.” *American Law And Economics Review.* 7 (2), 319–349.

——— and K. J. Gittings (2003). “Getting Off Death Row: Commuted Sentences And The Deterrent Effect Of Capital Punishment.” *Journal Of Law And Economics.* 46 (2), 453–478.

——— and T. G. Bali (2005). “Asymmetric Crime Cycles.” Working Paper No. W11210, National Bureau of Economic Research, Cambridge, Massachusetts.

———, S. C. Billups, and J. Overland (2005). “A Dynamic Model Of Differential Human Capital And Criminal Activity.” *Economica.* 72 (288), 655–681.

Moffett, M., A. K. Bohara, and K. Gawande (2005). “Governance And Performance: Theory-Based Evidence From U.S. Coast Guard Inspections.” *Policy Studies Journal.* 33 (2), 283–306.

Moffitt, T. E. (2005). “The New Look Of Behavioral Genetics In Developmental Psychopathology: Gene-Environment Interplay In Antisocial Behaviors.” *Psychological Bulletin.* 131 (4), 533–554.

Montmarquette, C., M. Nerlove, and P. Forest (1985). “Deterrence And Delinquency: An Analysis Of Individual Data.” *Journal Of Quantitative Criminology.* 1 (1), 37–58.

Morris, D. and L. Tweeten (1971). “The Cost Of Controlling Crime: A Study In Economies Of City Life.” *Annals Of Regional Science.* 5 (1), 33–49.

Mui, H. W. and M. M. Ali (1997). “Economic Analysis Of Crime And Punishment: An Asian Case.” *Applied Economics Letters.* 4 (4), 261–265.

Mullahy, J. and J. L. Sindelar (1994). “Do Drinkers Know When To Say When? An Empirical Analysis Of Drunk Driving.” *Economic Inquiry.* 32 (3), 383–394.

Muller, A. (1989). “Business Recession, Alcohol Consumption, Drinking And Driving Laws: Impact On Oklahoma Motor Vehicle Fatalities And Fatal Crashes.” *American Journal Of Public Health.* 79 (10), 1366–1370.

Müller, C. (2003). “Die Abschreckungshypothese im ökonomischen Modell der Kriminalität - Eine Meta-Analyse.” Master’s thesis, Darmstadt University Of Technology, Darmstadt.

Murphy, J. J., G. P. Allen, T. H. Stevens, and D. Weatherhead (2003). “A Meta-Analysis Of Hypothetical Bias In Stated Preference Valuation.” Working Paper 2003–8, University Of Massachusetts, Amherst.

Mustard, D. B. (2003). “Reexamining Criminal Behavior: The Importance Of Omitted Variable Bias.” *Review Of Economics And Statistics.* 85 (1), 205–211.

Myers, S. L. (1982). “Crime In Urban Areas: New Evidence And Results.” *Journal Of Urban Economics.* 11 (2), 148–158.

Nagel, W. G. (1977). “On Behalf Of A Moratorium On Prison Construction.” *Crime And Delinquency.* 23 (2), 154–172.

Nagin, D. S. (1978). “General Deterrence: A Review Of The Empirical Evidence.” in A. Blumstein, J. Cohen, and D. S. Nagin (eds.), *Deterrence And Incapacitation: Estimating The Effects Of Criminal Sanctions On Crime Rates*, National Academy Press, Washington D.C., pp. 95–139.

——— and G. Pogarsky (2001). “Integrating Celerity, Impulsivity, And Extralegal Sanctions Into A Model Of General Deterrence: Theory And Evidence.” *Criminology.* 39 (4), 865–892.

——— and G. Pogarsky (2003). “An Experimental Investigation Of Deterrence: Cheating, Self-Serving Bias, And Impulsivity.” *Criminology.* 41 (1), 167–193.

——— and R. Paternoster (1991). “The Preventive Effect Of The Perceived Risk Of Arrest: Testing An Expanded Conception Of Deterrence.” *Criminology.* 29 (4), 561–587.

Neumayer, E. (2003). “Good Policy Can Lower Violent Crime: Evidence From A Cross-National Panel Of Homicide Rates, 1980-97.” *Journal Of Peace Research.* 40 (6), 619–640.

Neustrom, M. W. and W. M. Norton (1993). “The Impact Of Drunk Driving Legislation In Louisiana.” *Journal Of Safety Research.* 24 (2), 107–121.

Nijkamp, P. and J. Poot (2002). “Meta-Analysis Of The Impact Of Fiscal Policies On Long-Run Growth.” Discussion Paper 02–028/3, Tinbergen Institute, Amsterdam.

——— and J. Poot (2005). "The Last Word On The Wage Curve?" *Journal Of Economic Surveys.* 19 (3), 421–450.

Nilsson, A. (2004). "Income Inequality And Crime: The Case Of Sweden." Working Paper 6–2004, Institute For Labour Market Policy Evaluation, Uppsala.

Niskanen, W. A. (1994). "Crime, Police, And Root Causes." Policy Analysis 218, Cato Institute, Washington D.C.

Norström, T. (1978). "Drunken Driving: A Tentative Causal Model." in R. Hauge (ed.), *Drinking-and-Driving in Scandinavia*, Vol. 6 of *Scandinavian Studies in Criminology*, The Scandinavian Research Council for Criminology, pp. 69–78.

——— (1983). "Law Enforcement And Alcohol Consumption Policy As Countermeasures Against Drunk Driving: Possibilities And Limitations." *Accident Analysis And Prevention.* 15 (6), 513–522.

Nott, D. J. and P. J. Green (2004). "Bayesian Variable Selection And The Swendsen-Wang Algorithm." *Journal Of Computational And Graphical Statistics.* 13 (1), 141–157.

Olson, D. E. (1997). "Testing Deterrence And Incapacitation As Crime Control Mechanisms: A Refinement Of The Hypothesis." PhD dissertation, University of Illinois, Chicago.

——— and M. D. Maltz (2001). "Right-To-Carry Concealed Weapon Laws And Homicide In Large U.S. Counties: The Effect On Weapon Types, Victim Characteristics, And Victim-Offender Relationships." *Journal Of Law And Economics.* 44 (2), 747–770.

Olson, M. K. (1999). "Agency, Rulemaking, Political Influences, Regulation, And Industry Compliance." *Journal Of Law, Economics And Organization.* 15 (3), 573–601.

Oosterbeek, H., R. Sloof, and G. van de Kuilen (2004). "Cultural Differences In Ultimatum Game Experiments: Evidence From A Meta-Analysis." *Experimental Economics.* 7 (2), 171–188.

Orviska, M. and J. Hudson (2003). "Tax Evasion, Civic Duty And The Law Abiding Citizen." *European Journal Of Political Economy.* 19 (1), 83–102.

Otterbein, K. F. (1979). "A Cross-Cultural Study Of Rape." *Aggressive Behavior.* 5 (4), 425–435.

Papachristos, A. V., T. L. Meares, and J. Fagan (2005). "Attention Felons: Evaluating Project Safe Neighborhoods In Chicago." Columbia Public Law Research Paper No. 05–97.

Papps, K. L. and R. Winkelmann (1999). "Unemployment And Crime: New Evidence To An Old Question." Discussion Paper 25, Updated, Institute For The Study Of Labor, Bonn.

Parilla, P. F. (1982). "Organizational Control Of Employee Theft: A Test Of Deterrence Hypothesis." PhD dissertation, University of Minnesota.

Park, C.-G. and J. K. Hyun (2003). "Examining The Determinants Of Tax Compliance By Experimental Data: A Case Of Korea." *Journal Of Policy Modeling.* 25 (8), 673–684.

Parker, J. S. and R. A. Atkins (1999). "Did The Corporate Criminal Sentencing Guidelines Matter? Some Preliminary Empirical Observations." *Journal Of Law And Economics.* 42 (1), 423–453.

Parker, R. N. (1995). "Bringing Booze Back In: The Relationship Between Alcohol And Homicide." *Journal Of Research In Crime And Delinquency.* 32 (1), 3–38.

——— and D. M. Smith (1979). "Deterrence, Poverty, And Type Of Homicide." *American Journal Of Sociology.* 85 (3), 614–624.

Parsley, T. J. (2001). "Basic Examination Of The Correlation Between Crime Rates And Income Inequality." Research Paper 2001–19, West Virginia University, Morgantown.

Passell, P. (1975). "The Deterrent Effect Of The Death Penalty: A Statistical Test." *Stanford Law Review.* 28 (1), 61–80.

——— and J. B. Taylor (1977). "The Deterrent Effect Of Capital Punishment: Another View." *American Economic Review.* 67 (3), 445–451.

Pastore, A. L. and K. Maguire (2004). *Sourcebook Of Criminal Justice Statistics.* http://www.albany.edu/sourcebook/, Accessed 2007.

Pate, A. M. and E. E. Hamilton (1992). "Formal And Informal Deterrents To Domestic Violence: The Dade County Spouse Assault Experiment." *American Sociological Review.* 57 (5), 691–697.

Paternoster, R. (1988). "Examining Three-Wave Deterrence Models: A Question On Temporal Order And Specification." *Journal Of Criminal Law And Criminology.* 79 (1), 135–179.

——— (1989a). "Absolute And Restrictive Deterrence In A Panel Of Youth: Explaining The Onset, Persistence/Desistance, And Frequency Of Delinquent Offending." *Social Problems.* 36 (3), 289–309.

——— (1989b). "Decisions To Participate In And Desist From Four Types Of Common Delinquency: Deterrence And The Rational Choice Perspective." *Law And Society Review.* 23 (1), 7–40.

——— and D. S. Nagin (1994). "Personal Capital And Social Control: The Deterrence Implications Of A Theory Of Individual Differences In Criminal Offending." *Criminology.* 32 (4), 581–606.

——— and L. Iovanni (1986). "The Deterrent Effect Of Perceived Severity: A Reexamination." *Social Forces.* 64 (3), 751–777.

——— and S. S. Simpson (1996). "Sanction Threats And Appeals To Morality: Testing A Rational Choice Model Of Corporate Crime." *Law And Society Review.* 30 (3), 549–583.

———, L. E. Saltzman, and G. P. Waldo (1983a). "Perceived Risk And Social Control: Do Sanctions Really Deter?" *Law And Society Review.* 17 (3), 457–479.

———, L. E. Saltzman, G. P. Waldo, and T. G. Chiricos (1982a). "Causal Ordering In Deterrence Research: An Examination Of Perceptions $\leftrightarrow$ Behavior Relationship." in J. Hagan (ed.), *Deterrence Reconsidered*, Beverly Hills, pp. 55–70.

———, L. E. Saltzman, G. P. Waldo, and T. G. Chiricos (1983b). "Estimating Perceptual Stability And Deterrent Effects: The Role Of Perceived Legal Punishment In The Inhibition Of Criminal Involvement." *Journal Of Criminal Law And Criminology.* 74 (1), 270–297.

———, L. E. Saltzman, T. G. Chiricos, and G. P. Waldo (1982b). "Perceived Risk And Deterrence: Methodological Artifacts In Perceptual Deterrence Research." *Journal Of Criminal Law And Criminology.* 73 (3), 1238–1258.

Pawlak, Z. (1982). “Rough Sets.” *International Journal Of Information And Computer Science.* 11 (5), 341–356.

——— (1991). *Rough Sets - Theoretical Aspects Of Reasoning About Data*, Kluwer Academic Publisher.

——— (1999). “Rough Classification.” *International Journal Of Human-Computer Studies.* 51 (2), 369–383.

Peck, J. K. (1976). “The Deterrent Effect Of Capital Punishment: Ehrlich And His Critics.” *Yale Law Journal.* 85 (3), 359–367.

Peek, C. W., H. P. Chalfant, and E. V. Milton (1979). “Sinners In The Hands Of An Angry God: Fundamentalist Fears About Drunken Driving.” *Journal For The Scientific Study Of Religion.* 18 (1), 29–39.

Pestello, H. F. (1984). “Deterrence: A Reconceptualization.” *Crime And Delinquency.* 30 (4), 593–609.

Petee, T. A., T. F. Milner, and M. R. Welch (1994). “Levels Of Social Integration In Group Contexts And The Effects Of Informal Sanction Threat On Deviance.” *Criminology.* 32 (1), 85–106.

Peterson, R. D. and W. C. Bailey (1991). “Felony Murder And Capital Punishment: An Examination Of The Deterrence Question.” *Criminology.* 29 (3), 367–395.

Pfeiffer, M. and C. Gelau (2002). “Determinanten regelkonformen Verhaltens am Beispiel des Straßenverkehrs. Variablen der Norminternalisierung im Zusammenwirken mit Effekten Polizeilicher Überwachungstätigkeit.” *Kölner Zeitschrift für Soziologie und Sozialpsychologie.* 54 (4), 694–713.

Philipson, T. J. and R. A. Posner (1996). “The Economic Epidemiology Of Crime.” *Journal Of Law And Economics.* 39 (2), 405–433.

Phillips, D. P. (1980). “The Deterrent Effect Of Capital Punishment: New Evidence On An Old Controversy.” *American Journal Of Sociology.* 86 (1), 139–148.

Phillips, L. and H. L. Votey (1975). “Crime Control In California.” *Journal Of Legal Studies.* 4 (2), 327–349.

Pierce, G. L. and W. J. Bowers (1981). “The Bartley-Fox Gun Law’s Short-Term Impact On Crime In Boston.” *Annals Of The American Academy Of Political And Social Science.* 455 (1), 120–132.

Piquero, A. R. and G. Pogarsky (2002). “Beyond Stafford And Warr’s Reconceptualization Of Deterrence: Personal And Vicarious Experiences, Impulsivity, And Offending Behavior.” *Journal Of Research In Crime And Delinquency.* 39 (2), 153–186.

——— and R. Paternoster (1998). “An Application Of Stafford And Warr’s Reconceptualization Of Deterrence To Drinking And Driving.” *Journal Of Research In Crime And Delinquency.* 35 (1), 3–39.

Plassmann, F. and T. N. Tideman (2001). “Does The Right To Carry Concealed Handguns Deter Countable Crimes? Only A Count Analysis Can Say.” *Journal Of Law And Economics.* 44 (2), 771–798.

Pogarsky, G. (2002). "Identifying Deterrable Offenders: Implications For Research On Deterrence." *Justice Quarterly.* 19 (3), 431–452.

——— (2004). "Projected Offending And Contemporaneous Rule-Violation: Implications For Heterotypic Continuity." *Criminology.* 42 (1), 111–135.

——— and A. R. Piquero (2004). "Studying The Reach Of Deterrence: Can Deterrence Theory Help Explain Police Misconduct?" *Journal Of Criminal Justice.* 32 (4), 371–386.

Pogue, T. F. (1975). "Effect Of Police Expenditures On Crime Rates: Some Evidence." *Public Finance Quarterly.* 3 (1), 14–44.

——— (1986). "Offender Expectations And Identification Of Crime Supply Functions." *Evaluation Review.* 10 (4), 455–482.

Polinsky, M. A. and S. M. Shavell (1999). "On The Disutility And Discounting Of Imprisonment And The Theory Of Deterrence." *Journal Of Legal Studies.* 28 (1), 1–16.

——— and S. M. Shavell (2000). "The Economic Theory Of The Public Enforcement Of Law." *Journal Of Economic Literature.* 38 (1), 45–76.

Pommerehne, W. W. and H. Weck-Hannemann (1996). "Tax Rates, Tax Administration And Income Tax Evasion In Switzerland." *Public Choice.* 88 (1–2), 161–170.

Pontell, H. N. (1978). "Deterrence: Theory Versus Practice." *Criminology.* 16 (1), 3–22.

——— (1979). "Deterrence And System Capacity: Crime And Punishment In California." PhD dissertation, State University of New York, Stony Brook.

Posada, C. E. (1994). "Modelos Económicos De La Criminalidad Y La Posibilidad De Una Dinámica Prolongada." *Plantatión Y Desarrollo.* 25 (edición especial), 217–225.

Pratt, T. C. (2004). "Assessing Macro-Level Predictors And Theories Of Crime: A Meta-Analysis." in M. Tonry (ed.), *Crime And Justice: A Review Of Research*, Vol. 32, University Of Chicago Press, Chicago, pp. 373–450.

Press, J. S. (1971). *Some Effects Of An Increase In Police Manpower In The 20th Precint Of New York City*, The Rand Corporation, New York.

Pudney, S., D. F. Deadman, and D. J. Pyle (2000a). "The Relationship Between Crime, Punishment And Economic Conditions: Is Reliable Inference Possible When Crimes Are Under-Recorded?" *Journal Of The Royal Statistical Society Series A - Statistics In Society.* 163 (1), 81–97.

———, D. J. Pyle, and T. Saruc (2000b). "Income Tax Evasion: An Experimental Approach." in Z. McDonald and D. Pyle (eds.), *Illicit Activity: The Economics Of Crime, Drugs And Tax Fraud*, Ashgate Publishers, Aldershot, pp. 267–289.

Pyle, D. J. and D. F. Deadman (1994). "Crime And The Business Cycle In Post-War Britain." *British Journal Of Criminology.* 34 (3), 339–357.

Quetelet, A. (1831). *Research On The Propensity For Crime At Different Ages*, Anderson, Cincinnati. Translated by Sawyer F. Sylvester, 1984.

R Development Core Team (2007). *R: A Language And Environment For Statistical Computing*, R Foundation For Statistical Computing. ISBN 3-900051-07-0.

Rabow, J., C. A. Neuman, R. K. Watts, and A. C. R. Hernandez (1987). "Alcohol-Related Hazardous Behavior Among College Students." in M. Galanter (ed.), *Recent Developments In Alcoholism*, Vol. 5, Plenum Press, New York, pp. 439–450.

Raftery, A. E., D. Madigan, and J. A. Hoeting (1997). "Bayesian Model Averaging For Linear Regression Models." *Journal Of The American Statistical Association.* 92 (437), 179–191.

Ralston, R. W. (1999). "Economy And Race: Interactive Determinants Of Property Crime In The United States, 1958–1995 - Reflections On The Supply Of Property Crime." *American Journal Of Economics And Sociology.* 58 (3), 405–434.

Rankin, J. H. and L. E. Wells (1982). "The Social Context Of Deterrence." *Sociology And Social Research.* 67 (1), 18–39.

Rasmussen, D. W., B. L. Benson, and D. L. Sollars (1993). "Spatial Competition In Illicit Drug Markets: The Consequences Of Increased Drug Law Enforcement." *Review Of Regional Studies.* 23 (3), 219–236.

Rawls, J. (1955). "Two Concept Of Rules." *Philosophical Review.* 64 (1), 3–32.

Reilly, B. and R. Witt (1996). "Crime, Deterrence And Unemployment In England And Wales: An Empirical Analysis." *Bulletin Of Economic Research.* 48 (2), 137–159.

Resignato, A. J. (2000). "Violent Crime: A Function Of Drug Use Or Drug Enforcement?" *Applied Economics.* 32 (6), 681–688.

Rettig, S. and H. E. Rawson (1963). "The Risk Hypothesis In Predictive Judgements Of Unethical Behavior." *Journal Of Abnormal And Social Psychology.* 66 (3), 243–248.

Retting, R. A., A. F. Williams, C. M. Farmer, and A. F. Feldman (1999). "Evaluation Of Red Light Camera Enforcement In Oxnard, California." *Accident Analysis And Prevention.* 31 (3), 169–174.

Rhee, L. and J. Zhang (1993). "Breath Testing In Canada: Deterrence Or Detection." *Applied Economics.* 25 (6), 765–775.

Richardson, L. E. (2003). "Deterring Drinking-And-Driving: State DWI Laws And Perceptions Of Punishment Costs." Working Paper, Harry S. Truman School Of Public Affairs, University Of Missouri, Columbia.

Ritsema, C. M., D. W. Thomas, and G. D. Ferrier (2003). "Economic And Behavioral Determinants Of Tax Compliance: Evidence From The 1997 Arkansas Tax Penalty Amnesty Program." IRS Publication 1500, IRS Statistics of Income Division.

Robertson, L. S., R. F. Rich, and L. H. Ross (1973). "Jail Sentences For Driving While Intoxicated In Chicago: A Judicial Policy That Failed." *Law And Society Review.* 8 (1), 55–67.

Rogers, P. N. and S. E. Schoenig (1994). "A Time Series Evaluation Of California's 1982 Driving-Under-The-Influence Legislative Reforms." *Accident Analysis And Prevention.* 26 (1), 63–78.

Rose, A. K. (2004). “A Meta-Analysis Of The Effect Of Common Currencies On International Trade.” Working Paper 10373, National Bureau of Economic Research, Cambridge, Massachusetts.

Rosenberg, M. S. (2005). “The File-Drawer Problem Revisited: A General Weighted Method For Calculating Fail-Safe Numbers In Meta-Analysis.” *Evolution.* 59 (2), 464–468.

Rosenthal, R. (1978). “Combining Results Of Independent Studies.” *Psychological Bulletin.* 85 (1), 185–193.

——— (1979). “The ’File Drawer Problem’ And Tolerance For Null Results.” *Psychological Bulletin.* 86 (3), 638–641.

Ross, H. L. (1975). “The Scandinavian Myth: The Effectiveness Of Drinking-And-Driving Legislation In Sweden And Norway.” *Journal Of Legal Studies.* 4 (2), 285–310.

——— (1977). “Deterrence Regained: The Cheshire Constabulary’s Breathalyser Blitz.” *Journal Of Legal Studies.* 6, 241–249.

——— (1987a). “Administrative License Revocation In New Mexico: An Evaluation.” *Law And Policy.* 9 (1), 5–16.

——— (1987b). “Britain’s Christmas Crusade Against Drinking And Driving.” *Journal Of Studies On Alcohol.* 48 (5), 476–482.

——— and H. Klette (1995). “Abandonment Of Mandatory Jail For Impaired Drivers In Norway And Sweden.” *Accident Analysis And Prevention.* 27 (2), 151–157.

——— and R. B. Voas (1990). “The New Philadelphia Story: The Effects Of Severe Punishment For Drunk Driving.” *Law And Policy.* 12 (1), 51–79.

———, D. T. Campbell, and G. V. Glass (1970). “Determining The Social Effects Of A Legal Reform: The British Breathalyser Crackdown Of 1967.” *American Behavioral Scientist.* 13 (4), 493–509.

———, R. McCleary, and G. LaFree (1990). “Can Mandatory Jail Laws Deter Drunk Driving? The Arizona Case.” *Journal Of Criminal Law And Criminology.* 81 (1), 156–170.

Rubinstein, A. (1980). “On An Anomaly Of The Deterrent Effect Of Punishment.” *Economics Letters.* 6 (1), 89–94.

Ruhm, C. J. (1996). “Alcohol Policies And Highway Vehicle Fatalities.” *Journal Of Health Economics.* 15 (4), 435–454.

Rupp, T. (2005). “Rough Set Methodology In Meta-Analysis: A Comparative And Exploratory Analysis.” Darmstadt Discussion Papers In Economics No. 157, Department Of Economics, Darmstadt.

——— (2006). “Multiple Measures In A Meta-Analysis: Guidelines In Special Cases.” Working Paper, Department Of Economics, Darmstadt.

Saffer, H. and F. J. Chaloupka (1989). “Breath Testing And Highway Fatality Rates.” *Applied Economics.* 21 (7), 901–912.

Sah, R. K. (1991). "Social Osmosis And Patterns Of Crime." *Journal Of Political Economy.* 99 (6), 1272–1295.

Sala-I-Martin, X. X. (1997). "I Just Ran Two Million Regressions." *American Economic Review.* 87 (2), 178–183.

Salem, R. G. and W. J. Bowers (1970). "Severity Of Formal Sanctions As A Deterrent To Deviant Behavior." *Law And Society Review.* 5 (1), 21–40.

Saltzman, L., R. Paternoster, and G. P. Waldo (1982). "Deterrent And Experiential Effects: The Problem Of Causal Order In Perceptual Deterrence Research." *Journal Of Research In Crime And Delinquency.* 19 (2), 172–189.

Sampson, R. J. and S. W. Raudenbusch (1999). "Systematic Observation Of Public Spaces: A New Look At Disorder In Urban Neighborhoods." *American Journal Of Sociology.* 105 (3), 603–651.

Sandelin, B. and G. Skogh (1986). "Property Crimes And The Police: An Empirical Analysis Of Swedish Data." *Scandinavian Journal Of Economics.* 88 (3), 547–561.

Scargle, J. D. (2000). "Publication Bias: The 'File-Drawer' Problem In Scientific Inference." *Journal Of Scientific Exploration.* 14 (1), 91–106.

Schuessler, K. F. (1952). "The Deterrent Influence Of The Death Penalty." *Annals Of American Academy Of Political And Social Science.* 284 (1), 54–62.

Schumann, K. F. and R. Kaulitzki (1991). "Limits Of General Deterrence: The Case Of Juvenile Delinquency." in K. Sessar and H.-J. Kerner (eds.), *Developments In Crime And Crime Control Research. German Studies On Victims, Offenders And The Public*, Springer, Berlin, pp. 1–21.

———, C. Berlitz, H.-W. Guth, and R. Kaulitzki (1987). *Jugendkriminalität und die Grenzen der Generalprävention*, Luchterhand, Köln.

Schwartz, B. (1984). *Psychology Of Learning And Behavior*, Norton, New York.

Scott, W. J. and H. G. Grasmick (1981). "Deterrence And Income Tax Cheating: Testing Interaction Hypotheses In Utilitarian Theories." *Journal Of Applied Behavioral Science.* 17 (3), 395–408.

Scribner, R. and D. Cohen (2001). "The Effect Of Enforcement On Merchant Compliance With The Minimum Legal Drinking Age Law." *Journal Of Drug Issues.* 31 (4), 857–866.

Sesnowitz, M. L. and D. McKee (1977). "Capital Punishment: The Canadian Experience." *Journal Of Behavioral Economics.* 6 (1–2), 145–154.

——— and L. J. Hexter (1982). "Economic Determinants Of Theft: Some Empirical Results." *Public Finance Quarterly.* 10 (4), 489–498.

Shachmurove, Y., G. F. Rengert, and S. Hakim (2001). "Target Search Of Burglars: A Revised Economic Model." *Papers In Regional Science.* 80 (2), 121–137.

Sharkansky, I. (1967). "Government Expenditures And Public Services In The American States." *American Political Science Review.* 61 (4), 1066–1077.

Shavit, Y. and A. Rattner (1988). "Age, Crime, And The Early Life Course." *American Journal Of Sociology.* 93 (6), 1457–1470.

Shaw, C. M. and H. D. McKay (1972). *Juvenile Delinquency And Urban Areas*, University Of Chicago Press, Chicago.

Sheffrin, S. M. and R. K. Triest (1992). "Can Brute Deterrence Backfire? Perceptions And Attitudes In Taxpayer Compliance." in J. Slemrod (ed.), *Why People Pay Taxes: Tax Compliance And Enforcement*, University Of Michigan Press, Ann Arbor, pp. 193–218.

Sheley, J. F. and K. D. Bailey (1985). "New Directions For Anti-Theft Policy: Reductions In Stolen Goods Buyers." *Journal Of Criminal Justice.* 13 (5), 399–415.

Shepard, E. M. and P. R. Blackley (2005). "Drug Enforcement And Crime: Recent Evidence From New York State." *Social Science Quarterly.* 86 (2), 323–342.

Shepherd, J. M. (2002a). "Fear Of The First Strike: The Full Deterrent Effect Of California's Two- And Three-Strikes Legislation." *Journal Of Legal Studies.* 31 (1), 159–201.

——— (2002b). "Police, Prosecutors, Criminals, And Determinate Sentencing: The Truth About Truth-In-Sentencing Laws." *Journal Of Law And Economics.* 45 (2), 509–534.

——— (2004). "Murders Of Passion, Execution Delays, And The Deterrence Of Capital Punishment." *Journal Of Legal Studies.* 33 (2), 283–322.

Sherman, L. W. and D. P. Rogan (1995). "Deterrent Effects Of Police Raids On Crack Houses: A Randomized, Controlled Experiment." *Justice Quarterly.* 12 (4), 755–781.

——— , D. Gottfredson, D. MacKenzie, P. Reuter, and S. Bushway (1998). "Preventing Crime: What Works, What Doesn't, What's Promising." National Institute Of Justice Report, U.S. Department Of Defense, Washington D.C.

Shimshack, J. P. and M. B. Ward (2005). "Regulator Reputation, Enforcement, And Environmental Compliance." *Journal Of Environmental Economics And Management.* 50 (3), 519–540.

Shore, E. R. and E. Maguin (1988). "Deterrence Of Drinking-Driving: The Effect Of Changes In The Kansas Driving Under The Influence Law." *Evaluation And Program Planning.* 11 (3), 245–254.

Sickles, R. C., P. Schmidt, and A. D. Witte (1979). "An Application Of The Simultaneous Tobit Model: A Study Of The Determinants Of Criminal Recidivism." *Journal Of Economics And Business.* 31 (3), 166–171.

Sigman, H. (1998). "Midnight Dumping: Public Policies And Illegal Disposal Of Used Oil." *Rand Journal Of Economics.* 29 (1), 157–178.

Silberman, M. (1976). "Toward A Theory Of Criminal Deterrence." *American Sociological Review.* 41 (3), 442–461.

Silverman, L. P. and N. L. Spruill (1977). "Urban Crime And The Price Of Heroin." *Journal Of Urban Economics.* 4 (1), 80–103.

Sinha, J. B. P. (1967). "Ethical Risk And Censure-Avoiding Behavior." *Journal Of Social Psychology.* 71 (2), 267–275.

——— and R. J. Wherry (1965). "Determinants Of Norm Violating Behaviour In A Simulated Industrial Setting." *Personnel Psychology.* 18 (4), 403–412.

Sirakaya, E. and M. Uysal (1997). "Can Sanctions And Rewards Explain Conformance Behaviour Of Tour Operators With Ecotourism Guidelines?" *Journal Of Sustainable Tourism.* 5 (4), 322–332.

Slemrod, J., M. Blumenthal, and C. W. Christian (2001). "Taxpayer Response To An Increased Probability Of Audit: Evidence From A Controlled Experiment In Minnesota." *Journal Of Public Economics.* 79 (3), 455–483.

Sloan, F. A. and P. B. Githens (1994). "Drinking, Driving, And The Price Of Automobile Insurance." *Journal Of Risk And Insurance.* 61 (1), 33–58.

———, B. A. Reilly, and C. Schenzler (1994). "Effects Of Prices, Civil And Criminal Sanctions, And Law Enforcement On Alcohol-Related Mortality." *Journal Of Studies On Alcohol.* 55 (4), 454–465.

———, B. A. Reilly, and C. Schenzler (1995). "Effects Of Tort Liability And Insurance On Heavy Drinking And Drinking And Driving." *Journal Of Law And Economics.* 38 (1), 49–77.

Smith, A. (1789). *An Inquiry Into The Nature And Causes Of The Wealth Of Nations* in 'The Harvard Classics.', P.F. Collier & Son, 1909-14, New York.

Smith, D. A. and R. Paternoster (1987). "The Gender Gap In Theories Of Deviance: Issues And Evidence." *Journal Of Research In Crime And Delinquency.* 24 (2), 140–172.

Smith, D. D., G. H. Givens, and R. L. Tweedie (1997). "Publication Bias In Meta-Analysis: A Bayesian Data-Augmentation Approach To Account For Issues Exemplified In The Passive Smoking Debate." *Statistical Science.* 12 (4), 221–250.

Smith, D. I. (1988). "Effect On Traffic Safety Of Introducing A 0.05% Blood Alcohol Level In Queensland, Australia." *Medicine, Science And The Law.* 28 (2), 165–170.

Smith, D. M., J. A. Devine, and J. F. Sheley (1992). "Crime And Unemployment: Effects Across Age And Race Categories." *Sociological Perspectives.* 35 (4), 551–572.

Smith, K. V. and J.-C. Huang (1995). "Can Markets Value Air Quality? A Meta-Analysis Of Hedonic Property Value Models." *Journal Of Political Economy.* 103 (1), 209–227.

Smith, K. W. (1992). "Reciprocity And Fairness: Positive Incentives For Tax Compliance." in J. Slemrod (ed.), *Why People Pay Taxes: Tax Compliance And Enforcement*, University Of Michigan Press, Ann Arbor, pp. 223–250.

Snortum, J. R. and D. E. Berger (1989). "Drinking-Driving Compliance In The United States: Perceptions And Behavior In 1983 And 1986." *Journal Of Studies On Alcohol.* 50 (4), 306–319.

Snyder, D. and C. Tilly (1972). "Hardship And Collective Violence In France, 1830 To 1960." *American Sociological Review.* 37 (5), 520–530.

Soares, R. R. (2004). "Development, Crime And Punishment: Accounting For The International Differences In Crime Rates." *Journal Of Development Economics.* 73 (1), 155–184.

Sollars, D. L., B. L. Benson, and D. W. Rasmussen (1994). "Drug Enforcement And The Deterrence Of Property Crime Among Local Jurisdictions." *Public Finance Quarterly.* 22 (1), 22–45.

Soper, J. R. and L. Thompson (1990). "Legislating Safety: A Look At The Effects Of Michigan's 1983 Anti-Drunk Driving Statute." *Policy Studies Journal.* 18 (4), 871–885.

Sorensen, J., R. Wrinkle, V. Brewer, and J. Marquart (1999). "Capital Punishment And Deterrence: Examining The Effect Of Executions On Murder In Texas." *Crime And Delinquency.* 45 (4), 481–493.

Spelman, W. (1993). "Abandoned Buildings: Magnets For Crime." *Journal Of Criminal Justice.* 21 (5), 481–495.

——— (2000). "The Limited Importance Of Prison Expansion." in A. Blumstein and J. Wallman (eds.), *The Crime Drop In America*, Cambridge University Press, New York, pp. 97–129.

Spencer, D. L. (1999). "Social Control, Delinquency, And Youth Status Achievement: A Developmental Approach." *Sociological Perspectives.* 42 (2), 305–324.

Spengler, H. (2004). "Ursachen und Kosten der Kriminalität in Deutschland - Drei empirische Untersuchungen." PhD dissertation, Darmstadt University Of Technology, Darmstadt.

Spicer, M. W. and E. J. Thomas (1982). "Audit Probabilities And The Tax Evasion Decision: An Experimental Approach." *Journal Of Economic Psychology.* 2 (3), 241–245.

——— and R. E. Hero (1985). "Tax Evasion And Heuristics: A Research Note." *Journal Of Public Economics.* 26 (2), 263–267.

——— and S. B. Lundstedt (1976). "Understanding Tax Evasion." *Public Finance.* 31 (2), 295–305.

Sridharan, S., S. Vujic, and S. J. Koopmann (2003). "Intervention Time Series Analysis Of Crime Rates." Discussion Paper 03–040/4, Tinbergen Institute, Free University Amsterdam, Amsterdam.

Stack, S. (1982). "Social Structure And Swedish Crime Rates: A Time-Series Analysis, 1950-1979." *Criminology.* 20 (3–4), 499–513.

——— (1984). "Income Inequality And Property Crime: A Cross-National Analysis Of Relative Deprivation Theory." *Criminology.* 22 (2), 229–257.

——— (1987). "Publicized Executions And Homicide, 1950-1980." *American Sociological Review.* 52 (4), 532–540.

——— (1990). "Execution Publicity And Homicide In South Carolina: A Research Note." *Sociological Quarterly.* 31 (4), 599–611.

Stafford, M. C., L. N. Gray, B. A. Menke, and D. A. Ward (1986). "Modeling The Deterrent Effects Of Punishment." *Social Psychology Quarterly.* 49 (4), 338–347.

Stafford, S. L. (2002). "The Effect Of Punishment On Firm Compliance With Hazardous Waste Regulations." *Journal Of Environmental Economics And Management.* 44 (2), 290–308.

Stalans, L. J., K. W. Smith, and K. A. Kinsey (1989). "When Do We Think About Detection? Structural Opportunity And Taxpaying Behavior." *Law And Social Inquiry.* 14 (3), 481–503.

Stanley, T. D. (2005a). “Beyond Publication Bias.” *Journal Of Economic Surveys.* 19 (3), 309–345.

——— (2005b). “Integrating The Empirical Tests Of The Natural Rate Hypothesis: A Meta-Regression Analysis.” *Kyklos.* 58 (4), 611–634.

——— and S. B. Jarell (2005). “Meta-Regression Analysis: A Quantitative Method Of Literature Surveys.” *Journal Of Economic Surveys.* 19 (3), 299–308.

Statacorp LP (2005). *Stata Statistical Software: Release 9.*

Steenbergen, M. R., K. M. McGraw, and J. T. Scholz (1992). “Taxpayer Adaption To The 1986 Tax Reform Act: Do New Tax Laws Affect The Way Taxpayers Think About Taxes?” in J. Slemrod (ed.), *Why People Pay Taxes: Tax Compliance And Enforcement*, University Of Michigan Press, Ann Arbor, pp. 9–38.

Steffensmeier, D. J. and M. D. Harer (1987). “Is The Crime Rate Really Falling: An ’Aging’ U.S. Population And Its Impact On The Nation’s Crime Rate, 1980-1984.” *Journal Of Research In Crime And Delinquency.* 24 (1), 23–48.

Sterling, T. D. (1959). “Publication Decisions And Their Possible Effects On Inference Drawn From Tests Of Significance - Or Vice Versa.” *Journal Of The American Statistical Association.* 54 (285), 30–34.

Stevans, L. K. (1988). “An Empirical Model Of Property Crime: Deterrence Versus Redistribution.” *Journal Of Post Keynesian Economics.* 10 (4), 572–584.

Stolzenberg, L. and S. J. D’Alessio (2000). “Gun Availability And Violent Crime: New Evidence From The National Incident-Based Reporting System.” *Social Forces.* 78 (4), 1461–1482.

——— and S. J. D’Alessio (2004). “Capital Punishment, Execution Publicity And Murder In Houston, Texas.” *Journal Of Criminal Law And Criminology.* 94 (2), 351–379.

Storey, D. J. (1979). “The Economics Of Environmental Law Enforcement Or Has The Prosecution Of Polluters Led To Cleaner Rivers In England And Wales?” *Environment And Planning A.* 11 (8), 897–918.

Stout, E., F. A. Sloan, and L. Liang (2000). “Reducing Harmful Alcohol-Related Behaviors: Effective Regulatory Methods.” *Journal Of Studies On Alcohol.* 61 (3), 402–412.

Sullivan, K., B. Keane, and C. Deffenti (2001). “Malingering On The RAVLT - Part I. Deterrence Strategies.” *Archives Of Clinical Neuropsychology.* 16 (7), 627–641.

Swan, P. L. (1984). “The Economics Of Law: Economic Imperialism In Negligence Law, No-Fault Insurance, Occupational Licensing And Criminology?” *Australian Economic Review.* 67 (3), 92–108.

Swimmer, E. (1974a). “Measurement Of The Effectiveness Of Urban Law Enforcement: A Simultaneous Approach.” *Southern Economic Journal.* 40 (4), 618–630.

——— (1974b). “The Relationship Of Police And Crime: Some Methodological And Empirical Results.” *Criminology.* 12 (3), 293–314.

Sykes, G. W. (1984). "Saturated Enforcement: The Efficacy Of Deterrence And Drunk Driving." *Journal Of Criminal Justice.* 12 (2), 185–197.

Tao, H.-L. (2004). "Property Crime Distribution And Equal Police Deployment: An Empirical Study Of Taiwan." *Journal Of Urban Economics.* 55 (1), 165–178.

Tauchen, H., A. D. Witte, and H. Griesinger (1994). "Criminal Deterrence: Revisiting The Issue With A Birth Cohort." *Review Of Economics And Statistics.* 76 (3), 399–412.

Taxman, F. S. and A. R. Piquero (1998). "On Preventing Drunk Driving Recidivism: An Examination Of Rehabilitation And Punishment Approaches." *Journal Of Criminal Justice.* 26 (2), 129–143.

Tay, R. (2005a). "Deterrent Effects Of Drinking Driving Enforcement: Some Evidence From New Zealand." *International Journal Of Transport Economics.* 32 (1), 103–109.

——— (2005b). "General And Specific Deterrent Effects Of Traffic Enforcement: Do We Have To Catch Offenders To Reduce Crashes?" *Journal Of Transport Economics And Policy.* 39 (2), 209–223.

Teevan, J. J. (1972). "Deterrent Effects Of Punishment: The Canadian Case." *Canadian Journal Of Criminology And Corrections.* 14 (1), 68–82.

——— (1976a). "Deterrent Effects Of Punishment For Breaking And Entering And Theft." in Law Reform Commision Of Canada (ed.), *Fear Of Punishment*, Ottawa.

——— (1976b). "Deterrent Effects Of Punishment: Subjective Measures Continued." *Canadian Journal Of Criminology And Corrections.* 18 (2), 152–160.

——— (1976c). "Subjective Perception Of Deterrence (Continued)." *Journal Of Research In Crime And Delinquency.* 13 (2), 155–164.

Thaler, R. H. (1977). "An Econometric Analysis Of Property Crime: Interaction Between Police And Criminals." *Journal Of Public Economics.* 8 (1), 37–51.

Theil, H. (1954). *Linear Aggregation Of Economic Relations*, North Holland Publishing Company, Amsterdam.

Thomson, E. (1997). "Deterrence Versus Brutalization: The Case Of Arizona." *Homicide Studies.* 1 (2), 110–128.

——— (1999). "Effects Of Execution On Homicides In California." *Homicide Studies.* 3 (2), 129–150.

Thornton, D., N. A. Gunningham, and R. A. Kagan (2005). "General Deterrence And Corporate Environmental Behavior." *Law And Policy.* 27 (2), 262–288.

Thurman, Q. C. (1991). "Taxpayer Noncompliance And General Prevention: An Expansion Of The Deterrence Model." *Public Finance.* 46 (2), 289–298.

Tibbetts, S. G. (1999). "Differences Between Women And Men Regarding Decisions To Commit Test Cheating." *Research In Higher Education.* 40 (3), 323–342.

Tittle, C. R. (1969). "Crime Rates And Legal Sanctions." *Social Problems.* 16 (4), 618–630.

——— (1974). “Sanction Fear And The Maintenance Of Social Order.” *Social Forces.* 55 (3), 579–596.

——— and A. R. Rowe (1973). “Moral Appeal, Sanction Threat, And Deviance: An Experimental Test.” *Social Problems.* 20 (4), 488–498.

——— and A. R. Rowe (1974). “Certainty Of Arrest And Crime Rates: A Further Test Of The Deterrence Hypothesis.” *Social Forces.* 52 (4), 455–462.

——— and E. V. Botchkovar (2005). “Self-Control, Criminal Motivation And Deterrence: An Investigation Using Russian Respondents.” *Criminology.* 43 (2), 307–353.

———, J. Bratton, and M. G. Gertz (2003). “A Test Of A Micro-Level Application Of Shaming Theory.” *Social Problems.* 50 (4), 592–617.

Trumbull, W. N. (1989). “Estimations Of The Economic Model Of Crime Using Aggregate And Individual Level Data.” *Southern Economic Journal.* 56 (2), 423–439.

Upper, J. R. and J. H. White (1976). “An Experimental Study Of General Deterrence: The Effect Of Threatened Punishment On Potential Break And Enter Offenders.” *Nederlands Tijdschrift Voor Criminologie.* 18, 68–81.

U.S. Supreme Court (1976). “Gregg V. Georgia.” Court Decision 428 U.S. 153, Supreme Court Of The United States. Chief Justice: Warren E. Burger; Associate Justices: William J. Brennan And Potter Stewart And Byron White And Thurgood Marshall And Harry Blackmun And Lewis Franklin Powell And William Rehnquist And John Paul Stevens.

Usher, D. (1997). “Education As A Deterrent To Crime.” *Canadian Journal Of Economics.* 30 (2), 367–384.

van der Sluis, J., M. van Praag, and W. Vijverberg (2003). “Entrepreneurship Selection And Performance.” Discussion Paper 03–046/3, Tinbergen Institute, Amsterdam.

van Tulder, F. (1992). “Crime, Detection Rate, And The Police: A Macro Approach.” *Journal Of Quantitative Criminology.* 8 (1), 113–131.

——— and A. van der Torre (1999). “Modeling Crime And The Law Enforcement System.” *International Review Of Law And Economics.* 19 (4), 471–486.

Vandaele, W. (1978a). “An Econometric Model Of Auto Theft In The United States.” in J. M. Heineke (ed.), *Economic Models Of Criminal Behavior*, North Holland Publishing Company, New York, pp. 303–390.

——— (1978b). “Participation In Illegitimate Activities: Ehrlich Revisited.” in A. Blumstein, J. Cohen, and D. S. Nagin (eds.), *Deterrence And Incapacitation: Estimating The Effects Of Criminal Sanctions On Crime Rates*, National Academy Press, Washington D.C., pp. 270–335.

Velez, L. F., V. E. Espitia, H. Banguero, F. Mendez, E. Muñoz, W. Rotawinsky, G. Vanegas, and R. Espinoza (1999). “Victimization In Colombia: The City Of Cali: An Exploratory Analysis.” World Bank Project On Crime In LAC Cities, Cali.

Venables, W. N. and B. D. Ripley (2002). *Modern Applied Statistics With S*, fourth edition ed., Springer, Berlin.

Victor, M. I. (1977). “Relations Between Known Crime And Police Spending In Large United States Cities.” *Sociological Focus.* 10 (2), 199–206.

Vingilis, E. R., H. Blefgen, H. Lei, K. Sykora, and R. E. Mann (1988). “An Evaluation Of The Deterrent Impact Of Ontario’s 12-Hour Licence Suspension Law.” *Accident Analysis And Prevention.* 20 (1), 9–17.

Vinod, H. D. (1999). “Statistical Analysis Of Corruption Data And Using The Internet To Reduce Corruption.” *Journal Of Asian Economics.* 10 (4), 591–603.

Virén, M. (1990). “A Note On Finnish Property Criminality.” *International Journal Of Social Economics.* 17 (9), 55–59.

——— (1994). “A Test Of An Economics Of Crime Model.” *International Review Of Law And Economics.* 14 (3), 363–370.

——— (2001). “Modelling Crime And Punishment.” *Applied Economics.* 33 (14), 1869–1879.

Viscusi, K. W. (1986). “The Risks And Rewards Of Criminal Activity: A Comprehensive Test Of Criminal Deterrence.” *Journal Of Labor Economics.* 4 (3), 317–340.

——— (2000). “The Value Of Life In Legal Contexts: Survey And Critique.” *American Law And Economics Review.* 2 (1), 195–210.

Voas, R. B. (1986). “Evaluation Of Jail As A Penalty For Drunk Driving.” *Alcohol, Drugs And Driving.* 2 (2), 47–70.

———, A. S. Tippetts, and J. E. Lange (1997). “Evaluation Of A Method For Reducing Unlicensed Driving: The Washington And Oregon License Plate Sticker Laws.” *Accident Analysis And Prevention.* 29 (5), 627–634.

——— and J. M. Hause (1987). “Deterring The Drinking Driver: The Stockton Experience.” *Accident Analysis And Prevention.* 19 (2), 81–90.

von Hofer, H. and H. Tham (1989). “General Deterrence In A Longitudinal Perspective. A Swedish Case: Theft, 1841–1985.” *European Sociological Review.* 5 (1), 25–45.

von Mayr, G. (1867). *Statistik der gerichtlichen Polizei im Königreiche Bayern und in einigen anderen Ländern*, Statistics Bavaria, Munich.

Votey, H. L. (1978). “The Deterrence Of Drunken Driving In Norway And Sweden: An Econometric Analysis Of Existing Policies.” in R. Hauge (ed.), *Drinking-And-Driving In Scandinavia*, Universitetsforlaget, Oslo, pp. 79–99.

——— (1984a). “Crime In Sweden: Causality, Control Effects, And Economic Efficiency.” in K. Smith and A. Witte (eds.), *Advances In Applied Micro Economics*, Vol. 3, JAI Press, Greenwich, pp. 229–265.

——— (1984b). “Recent Evidence From Scandinavia On Deterring Alcohol Impaired Driving.” *Accident Analysis And Prevention.* 16 (2), 123–138.

——— and P. Shapiro (1983). “Highway Accidents In Sweden: Modelling The Process Of Drunken Driving Behaviour And Control.” *Accident Analysis And Prevention.* 15 (6), 523–533.

Wagenaar, A. C., T. L. Toomey, and D. J. Erickson (2005). “Preventing Youth Access To Alcohol: Outcomes From A Multi-Community Time-Series Trial.” *Addiction.* 100 (3), 335–345.

Waldo, G. P. and T. G. Chiricos (1972). “Perceived Penal Sanction And Self-Reported Criminality: Neglected Approach To Deterrence Research.” *Social Problems.* 19 (4), 522–540.

Waldorf, B. and P. Byun (2005). “Meta-Analysis Of The Impact Of Age Structure On Fertility.” *Journal Of Population Economics.* 18 (1), 15–40.

Ward, D. A., B. A. Menke, L. N. Gray, and M. C. Stafford (1986). “Sanctions, Modeling, And Deviant Behavior.” *Journal Of Criminal Justice.* 14 (6), 501–508.

Ward, S., R. Bachmann, and R. Paternoster (1992). “The Rationality Of Sexual Offending: Testing A Deterrence/Rational Choice Conception Of Sexual Assault.” *Law And Society Review.* 26 (2), 343–372.

Wärneryd, K. E. and B. Walerud (1982). “Taxes And Economic Behavior: Some Interview Data On Tax Evasion In Sweden.” *Journal Of Economic Psychology.* 2 (3), 187–211.

Watson, R. E. L. (1986). “The Effectiveness Of Increased Police Enforcement As A General Deterrent.” *Law And Society Review.* 20 (2), 293–299.

Weber, J. M. and R. E. Crew (2000). “Deterrence Theory And Marine Oil Spills: Do Coast Guard Civil Penalties Deter Pollution?” *Journal Of Environmental Management.* 58 (3), 161–168.

Webley, P. and S. Halstead (1986). “Tax Evasion On The Micro: Significant Simulations Or Expedient Experiments.” *Journal Of Interdisciplinary Economics.* 1 (1), 87–100.

Weichselbaumer, D. and R. Winter-Ebmer (2005). “A Meta-Analysis Of The International Gender Wage Gap.” *Journal Of Economic Surveys.* 19 (3), 479–511.

Weinberg, B., E. Gould, and D. B. Mustard (2002). “Crime Rates And Local Labor Market Opportunities In The United States: 1979–1997.” *Review Of Economics And Statistics.* 84 (1), 45–61.

Welch, M., Y. Xu, T. Bjarnason, T. Petee, P. O’Donnell, and P. Magro (2005). “But Everybody Does It...: The Effects Of Perceptions, Moral Pressures, And Informal Sanctions On Tax Cheating.” *Sociological Spectrum.* 25 (1), 21–52.

Wellford, C. R. (1974). “Crime And The Police: A Multivariate Analysis.” *Criminology.* 12 (2), 195–213.

Wenzel, M. (2004). “The Social Side Of Sanctions: Personal And Social Norms As Moderators Of Deterrence.” *Law And Human Behavior.* 28 (5), 547–567.

West, A. D. (2002). “Death As Deterrent Or Prosecutorial Tool? Examining The Impact Of Louisiana’s Child Rape Law.” *Criminal Justice Policy Review.* 13 (2), 156–191.

Whetten-Goldstein, K., F. A. Sloan, E. Stout, and L. Liang (2000). “Civil Liability, Criminal Law, And Other Policies And Alcohol Related Motor Vehicle Fatalities In The United States, 1984–1995.” *Accident Analysis And Prevention.* 32 (6), 723–733.

Wieczorek, W. F., A. L. Mirand, and C. P. A. Callahan (1994). “Perception Of The Risk Of Arrest For Drinking And Driving.” *Criminal Justice And Behavior.* 21 (3), 312–324.

Wilkinson, J. T. (1987). "Reducing Drunken Driving: Which Policies Are Most Effective." *Southern Economic Journal.* 54 (2), 322–334.

Williams, A. F. and L. S. Robertson (1975). "The Fatal Crash Reduction Program: A Reevaluation." *Accident Analysis And Prevention.* 7 (1), 37–47.

Williams, F. P. (1985). "Deterrence And Social Control: Rethinking The Relationship." *Journal Of Criminal Justice.* 13 (2), 141–151.

Williams, J. and R. C. Sickles (2002). "An Analysis Of The Crime As Work Model: Evidence From The 1958 Philadelphia Birth Cohort Study." *Journal Of Human Resources.* 37 (3), 479–509.

Williams, K. R. (1992). "Social Sources Of Marital Violence And Deterrence: Testing An Integrated Theory Of Assaults Between Partners." *Journal Of Marriage And The Family.* 54 (3), 620–629.

——— and J. P. Gibbs (1981). "Deterrence And Knowledge Of Statutory Penalties." *Sociological Quarterly.* 22 (4), 591–606.

Willis, K. G. (1983). "Spatial Variations In Crime In England And Wales: Testing An Economic Model." *Regional Studies.* 17 (4), 261–272.

Wilson, D. P. (2005). "Additional Law Enforcement As A Deterrent To Criminal Behavior: Empirical Evidence From The National Hockey League." *Journal Of Socio-Economics.* 34 (3), 319–330.

Wilson, J. L. and S. M. Sheffrin (2005). "Understanding Surveys Of Taxpayer Honesty." *Finanz-Archiv.* 61 (2), 256–274.

Wilson, J. Q. and B. Boland (1978). "The Effect Of The Police On Crime." *Law And Society Review.* 12 (3), 367–390.

——— and G. L. Kelling (1982). "Broken Windows: The Police And Neighborhood Safety." *Atlantic Monthly.* 249 (3), 29–38.

Wilson, R. J. and B. A. Jonah (1985). "Identifying Impaired Drivers Among The General Driving Population." *Journal Of Studies On Alcohol.* 46 (6), 531–537.

Winter-Ebmer, R. and S. Raphael (2001). "Identifying The Effect Of Unemployment On Crime." *Journal Of Law And Economics.* 44 (1), 259–283.

Withers, G. (1984). "Crime, Punishment And Deterrence In Australia: An Empirical Investigation." *Economic Record.* 60 (169), 176–185.

Witt, R. (2005). "Do Players React To Anticipated Sanction Changes? Evidence From The English Premier League." *Scottish Journal Of Political Economy.* 52 (4), 623–640.

———, A. Clarke, and N. Fielding (1998). "Crime, Earnings Inequality And Unemployment In England And Wales." *Applied Economics Letters.* 5 (4), 265–267.

———, A. Clarke, and N. Fielding (1999). "Crime And Economic Activity: A Panel Data Approach." *British Journal Of Criminology.* 39 (3), 391–400.

——— and A. D. Witte (2000). "Crime, Prison, And Female Labor Supply." *Journal Of Quantitative Criminology.* 16 (1), 69–85.

Witte, A. D. (1980). "Estimating The Economic Model Of Crime With Individual Data." *Quarterly Journal Of Economics.* 94 (1), 57–84.

——— and D. F. Woodbury (1985). "The Effect Of Tax Laws And Tax Administration On Tax Compliance: The Case Of The U.S. Individual Income Tax." *National Tax Journal.* 38 (1), 1–14.

——— and P. Schmidt (1977). "An Analysis Of Recidivism, Using The Truncated Lognormal Distribution." *Applied Statistics.* 26 (3), 302–311.

——— and R. Witt (2001). "What We Spend And What We Get: Public And Private Provision Of Crime Prevention." *Fiscal Studies.* 22 (1), 1–40.

Wolpin, K. I. (1978a). "Capital Punishment And Homicide In England: A Summary Of Results." *American Economic Review.* 68 (2), 422–427.

——— (1978b). "An Economic Analysis Of Crime And Punishment In England And Wales, 1894–1967." *Journal Of Political Economy.* 86 (5), 815–840.

——— (1980). "A Time Series-Cross Section Analysis Of International Variation In Crime And Punishment." *Review Of Economics And Statistics.* 62 (3), 417–423.

Wong, Y.-C. R. (1995). "An Economic Analysis Of The Crime Rate In England And Wales, 1857–92." *Economica.* 62 (246), 235–246.

Worrall, J. L. and T. C. Pratt (2004). "On The Consequences Of Ignoring Unobserved Heterogeneity When Estimating Macro-Level Models Of Crime." *Social Science Research.* 33 (1), 79–105.

Wright, B. R., A. Caspi, T. E. Moffitt, and R. Paternoster (2004). "Does The Perceived Risk Of Punishment Deter Criminally Prone Individuals? Rational Choice, Self-Control, And Crime." *Journal Of Research In Crime And Delinquency.* 41 (2), 180–213.

Yu, J. (2000). "Punishment And Alcohol Problems: Recidivism Among Drinking-Driving Offenders." *Journal Of Criminal Justice.* 28 (4), 261–270.

——— and A. E. Liska (1993). "The Certainty Of Punishment: A Reference Group Effect And Its Functional Form." *Criminology.* 31 (3), 447–464.

Yunker, J. A. (1976). "Is The Death Penalty A Deterrent To Homicide? Some Time Series Evidence." *Journal Of Behavioral Economics.* 5 (1), 45–81.

——— (2001). "A New Statistical Analysis Of Capital Punishment Incorporating U.S. Postmoratorium Data." *Social Science Quarterly.* 82 (2), 297–311.

Zador, P. L. (1976). "Statistical Evaluation Of The Effectiveness Of Alcohol Safety Action Projects." *Accident Analysis And Prevention.* 8 (1), 51–66.

———, A. K. Lund, M. Fields, and K. Weinberg (1989). "Fatal Crash Involvement And Laws Against Alcohol-Impaired Driving." *Journal Of Public Health Policy.* 10 (4), 467–485.

Zedlewski, E. W. (1983). “Deterrence Findings And Data Sources: A Comparison Of The Uniform Crime Reports And The National Crime Surveys.” *Journal Of Research In Crime And Delinquency.* 20 (2), 262–276.

Zhang, J. (1997). “The Effect Of Welfare Programs On Criminal Behavior: A Theoretical And Empirical Analysis.” *Economic Inquiry.* 35 (1), 120–137.

Zhiqiang, L. (2004). “Capital Punishment And The Deterrence Hypothesis: Some New Insights And Empirical Evidence.” *Eastern Economic Journal.* 30 (2), 237–258.

Zimmerman, P. R. (2004). “State Executions, Deterrence, And The Incidence Of Murder.” *Journal Of Applied Economics.* 7 (1), 163–193.

Zimring, F. E. (1984). “Youth Homicide In New York: A Preliminary Analysis.” *Journal Of Legal Studies.* 13 (1), 81–99.

——— and G. J. Hawkins (1973). *Deterrence: The Legal Threat In Crime Control*, University Of Chicago Press, Chicago.

CURRICULUM VITAE

Thomas Rupp

Adresse:

Büro:	Privat:
Technische Universität Darmstadt	Ringstr. 51
Fachgebiet Empirische Wirtschaftsforschung und Mikroökonometrie	D-63533 Mainhausen
Residenzschloss, Marktplatz 15	
D-64283 Darmstadt	

Persönliche Daten:

Geboren am 26.06.1977 in Montevideo (Uruguay), ledig, keine Kinder
Deutscher Staatsbürger

Bildungsweg:

Dissertation	Abschluss des Promotionsverfahrens zum Dr. rer. pol. mit Auszeichnung, TU Darmstadt, 2008
Promotion	Promotionsstudium im Fachgebiet Empirische Wirtschaftsforschung und Mikroökonometrie, TU Darmstadt, SS 2004 bis SS 2008
Diplom	Abschluss als Dipl.-Math. mit der Note "sehr gut", Universität Frankfurt, 2003
Studium	Mathematik mit Nebenfach Volkswirtschaftslehre (und Informatik) an der Universität Frankfurt, WS 1998/1999 bis SS 2003
Abitur	Kreuzburggymnasium Großkrotzenburg, 1997

Berufserfahrung:

seit 1.2007	Wissenschaftlicher Mitarbeiter am Institut für Volkswirtschaftslehre des Fachbereichs Rechts- und Wirtschaftswissenschaften am Fachgebiet Empirische Wirtschaftsforschung und Mikroökonometrie der TU Darmstadt
seit 2.2004	Betreuung und Organisation von Seminaren, Studien- und Diplomarbeiten

2.2004 bis 1.2007	Wissenschaftlicher Mitarbeiter im DFG-Projekt “Meta-Analyse empirischer Abschreckungsstudien - ein quantitiver methodenkritischer Vergleich kriminologischer und ökonomischer Untersuchungen zur negativen Generalprävention” an der TU Darmstadt
2000 bis 2003	Tutor und Übungsleiter (als wiss. Hilfskraft) von sechs Lehrveranstaltungen an der Universität Frankfurt (“Mikroökonomie I”, “Statistik”, “Stochastik für Informatiker” und drei Mal “stochastic processes”)
8.1997 bis 8.1998	Zivildienst in der Caritas Sozialstation St. Laurentius e.V. Kahl/Karlstein
7.1997 bis 8.1997	Arbeit in diversen Funktionen in der Caritas Sozialstation St. Laurentius e.V. Kahl/Karlstein
10.1996 bis 5.1998	Systembetreuer im Beratungs- und Softwarehaus **fecher**.

Sonstige universitäres Engagement:

2007 bis 2008	Mitglied der Berufungskommission zur Besetzung der Professur “Empirische Wirtschaftsforschung” im Fachbereich Rechts- und Wirtschaftswissenschaften
2002 bis 2003	Studentischer Vertreter im neu gegründeten “Institut für Stochastik und mathematische Informatik” (ISMI) an der Universität Frankfurt
1999 bis 2003	Stipendium der “Jugendstiftung”

Sonstiges:

Kenntnisse	Unter anderem in: diverse Windows- und Linux Plattformen, Java, PhP, HTML, LaTeX, R, Stata, Maple
Hobbies	Informatik und die Geschichte des Mittelalters, Computer, Fußball, Fahrrad, tabletop Rollenspiel, Puzzles

Publikationen und Vorträge:

Publikationen:

Metaanalyse empirischer Abschreckungsstudien - Untersuchungsansatz und erste empirische Befunde, in Lösel/Bender/Jehle (Hrsg.): Kriminologie und wissensbasierte Kriminalpolitik, Entwicklungs- und Evaluationsforschung; Neue Kriminologische Schriftenreihe, Forum Verlag Godesberg (2007) S. 632-648 mit D. Dölling, H. Entorf, D. Hermann und A. Woll)

Zur generalpräventiven Abschreckungswirkung des Strafrechts - Befunde einer Metaanalyse, Soziale Probleme, Sonderausgabe: Härtere Strafen - weniger Kriminalität? Zur Verschärfung der Sanktionseinstellungen (2007) mit D. Dölling, D. Hermann, A. Häring und A. Woll

Diskussionsbeiträge:
Meta Analysis of Empirical Deterrence Studies: an explorative contest, Darmstadt Discussion Papers in Economics Nr. 174, Darmstadt (2006)

Rational Actors in Balancing Markets: a Game-Theoretic Model and Results, Darmstadt Discussion Papers in Economics Nr. 171, Darmstadt (2006)

Metaanalyse empirischer Abschreckungsstudien - Untersuchungsansatz und erste empirische Befunde, Darmstadt Discussion Papers in Economics Nr. 170, Darmstadt (2006) mit D. Dölling, H. Entorf, D. Hermann und A. Woll

Rough set methodology in meta-analysis - a comparative and exploratory analysis, Darmstadt Discussion Papers in Economics Nr. 157, Darmstadt (2005)

Vorträge:
Rational Actors and Balancing Markets - A Game-Theoretic Model, Jahrestagung 2007 des Vereins für Socialpolitik, München (9.10.-12.10.2007)

Meta-Analysis of Empirical Deterrence Studies - Basic Concepts and General Results, The Stockholm Criminology Symposium (4.6.-6.6.2007)

Publication Bias in Criminometrics - Motivation, Problems and Preliminary Results, Jahrestagung 2006 des Vereins für Socialpolitik, Bayreuth (26.9.-29.9.2006)

Meta-Analysis of Empirical Deterrence Studies - Basic Concepts and Preliminary Results, 6th Annual Conference of the European Society of Criminology, Tübingen (26.8.-29.8.2006)

Meta-Analyse empirischer Abschreckungsstudien - Untersuchungsansatz und erste Befunde, 42. kriminologisches Kolloquium der südwestdeutschen und schweizerischen kriminologischen Institute, Flehingen (30.6.-2.7.2006)