JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS Vol. 44, No. 4, Aug. 2009, pp. 883–909

doi:10.1017/S0022109009990159

Is There an Intertemporal Relation between Downside Risk and Expected Returns?

Turan G. Bâli, K. Ozgur Demirtas, and Haim Levy*

Abstract

This paper examines the intertemporal relation between downside risk and expected stock returns. Value at Risk (VaR), expected shortfall, and tail risk are used as measures of downside risk to determine the existence and significance of a risk-return tradeoff. We find a positive and significant relation between downside risk and the portfolio returns on NYSE/AMEX/Nasdaq stocks. VaR remains a superior measure of risk when compared with the traditional risk measures. These results are robust across different stock market indices, different measures of downside risk, loss probability levels, and after controlling for macroeconomic variables and volatility over different holding periods as originally proposed by Harrison and Zhang (1999).

I. Introduction

The conditional mean and variance of return on the market portfolio play central roles in Merton's (1973) intertemporal capital asset pricing model (ICAPM). Although theoretical models suggest a positive relation between risk and return for the aggregate stock market, the existing empirical literature fails to agree on the intertemporal relation between expected return and volatility. There is a long literature that has tried to identify the existence of such a tradeoff between risk and return, but the results are far from being conclusive (see the recent article by Guo and Whitelaw (2006) and the references therein).

*Bali, turan.bali@baruch.cuny.edu, Demirtas, ozgur.demirtas@baruch.cuny.edu, Zicklin School of Business, Baruch College, City University of New York, One Bernard Baruch Way, Box 10-225, New York, NY 10010; and Levy, mshlevy@mscc.huji.ac.il, Jerusalem School of Business Administration, Hebrew University of Jerusalem, 91905, Israel. We thank Hendrik Bessembinder (the editor) and Joseph Chen (the referee) for their extremely helpful comments and suggestions. We also thank George Aragon, Pierluigi Balduzzi, Jennifer Carpenter, Wayne Ferson, Stephen Figlewski, Campbell Harvey, Salih Neftci, Maureen O'Hara, Robert Schwartz, and Robert Whitelaw. An earlier version of this paper was presented at Baruch College, Cornell University, New York University, and the Graduate School and University Center of the City University of New York. Bali gratefully acknowledges the financial support of the PSC-CUNY Research Foundation of the City University of New York. Demirtas gratefully acknowledges the financial support of the PSC-CUNY Research Foundation of the City University of New York and of the Eugene M. Lang Research Foundation. Levy acknowledges the financial support of the Krueger Center of Finance. All errors remain our responsibility.

This paper examines the intertemporal relation between downside risk and expected return on the market. Value at Risk (VaR), expected shortfall (ES), and tail risk (TR) are used as measures of downside risk to determine the existence and significance of a risk-return tradeoff for several stock market indices. The results indicate a positive and significant relation between VaR and the value- and equal-weighted portfolio returns on NYSE/AMEX/Nasdaq stocks. This finding also holds for the NYSE/AMEX, NYSE, Nasdaq, and S&P 500 index portfolios. As alternative measures of downside risk, we also consider ES and TR, which measure the mean and variance of losses beyond some VaR thresholds, respectively. We show that the strong positive relation between downside risk and excess market return is robust across different left-tail risk measures.[1]

Ghysels, Santa-Clara, and Valkanov (2005) have recently shown that the choice of window size (from 1 to 6 months) in the estimation of realized variance has tremendous impact on the significance of risk-return tradeoff.[2] When they compute the realized variance as the sum of squared daily returns over the past 1 month, they find no evidence of a significant link between realized variance and future market returns. However, they report a significantly positive relation between the excess market return and the realized variance obtained from the past 3 to 6 months of daily data. In this paper, we compare the relative performance of various VaRs and realized variance measures computed over different horizons in predictive regressions. VaR remains a superior measure of risk even when compared to the traditional risk measures that have significant predictive power for market returns. These results are robust across different loss probability levels and after controlling for macroeconomic variables associated with business cycle fluctuations.

There are several reasons why we consider downside risk in determining the existence of a positive risk-return tradeoff. First, there is a long literature about safety-first investors who minimize the chance of disaster (or the probability of failure). The portfolio choice of a safety-first investor is intended to maximize expected return, subject to a downside risk constraint. The safety-first investor of Roy (1952), Baumol (1963), Levy and Sarnat (1972), and Arzac and Bawa (1977) uses a downside risk measure that is a function of VaR. Roy (1952) indicates that most investors are principally concerned with avoiding a possible disaster and

[1]Ang, Chen, and Xing (2006) decompose market beta into down versus up beta and focus on the cross-sectional predictive power of down beta (which they term downside risk). Their firm-level Fama-MacBeth (1973) regressions and long-short portfolio analyses indicate that there is a positive and significant relation between down beta and the cross-section of expected returns, whereas up beta cannot explain the cross-sectional variation in average stock returns. We should note that in Ang et al. (2006) downside risk is defined as a systematic risk conditional on market declines, whereas in our paper downside risk is defined as an extreme measure of risk such as VaR, ES, or TR. We focus on the time-series relation between extreme risk measures and expected returns on the aggregate market portfolio, whereas Ang et al. (2006) examine the cross-sectional relation between down beta and individual stock returns.

[2]Note that Harrison and Zhang (1999) are the first to look at the time-series relation between expected returns and conditional volatility over different holding periods and across different states of the economy and are the first to uncover a significantly positive risk and return relation at long holding intervals.

that the principle of safety plays a crucial role in the decision-making process. Roy ((1952), p. 432) states:

> Decisions taken in practice are less concerned with whether a little more of this or of that will yield the largest net increase in satisfaction than with avoiding known rocks of uncertain position or with deploying forces so that, if there is an ambush round the next corner, total disaster is avoided. If economic survival is always taken for granted, the rules of behavior applicable in an uncertain and ruthless world cannot be covered.

Thus, the idea of a disaster exists, and a risk averse safety-first investor will seek to reduce the chance of such a catastrophe occurring as much as possible.

Second, commercial banks, investment banks, insurance companies, and nonfinancial firms hold portfolios of assets that may include stocks, bonds, currencies, and derivatives. Each institution needs to quantify the amount of risk its portfolio may incur in the course of a day, week, month, or year. For example, a bank needs to assess its potential losses in order to put aside enough capital to cover them. Similarly, a company needs to track the value of its assets and any cash flows resulting from losses on its portfolio. In addition, credit rating and regulatory agencies must be able to assess likely losses on portfolios as well, since they need to set capital requirements and issue credit ratings. These institutions can judge the likelihood and magnitude of potential losses on their portfolios using VaR. Regulatory concerns require commercial banks to report a single number, the so-called VaR, which measures the maximum loss on their trading portfolio if the lowest 1% quantile return were to materialize. Capital adequacy is judged on the basis of the size of this expected loss. Likewise, pension funds are often required by law to structure their investment portfolio such that the risk of underfunding is kept low (e.g., equity investment may be capped).

Third, asset returns have been modeled in continuous time as diffusions by Black and Scholes (1973), as pure jump processes by Cox and Ross (1976), and as jump diffusions by Merton (1976). The rationale usually given for describing asset returns as jump diffusions is that diffusions capture frequent small moves, while jumps capture rare large moves. Carr, Geman, Madan, and Yor (2002) develop a continuous time model that allows for both diffusions and jumps of either finite or infinite activity. They find that market index returns tend to be pure jump processes of infinite activity and finite variation, and thus the index return processes appear to have effectively diversified away any diffusion risk. They indicate that the jump components account for significant skewness levels that statistically may be either positive or negative but that are risk-neutrally negative. They report significantly greater skewness and kurtosis in the risk-neutral process than in the statistical process. The results presented in Carr et al. (2002) suggest that extreme movements in stock returns can be interpreted as signals, whereas the frequent small fluctuations can be viewed as noise that may not have the power to explain time-series variation in excess market returns.

Finally, the mean-variance analysis developed by Markowitz (1952), (1959) relies critically on two assumptions: Either the investors have a quadratic utility or the asset returns are jointly normally distributed (see Levy and Markowitz (1979), Chamberlain (1983), and Berk (1997)). Both assumptions are not required, just one or the other: i) If an investor has quadratic preferences, she cares only about the mean and variance of returns, but she will not care about extreme losses. ii) Mean-variance optimization can be justified if the asset returns are jointly normally distributed, since the mean and variance will completely describe the distribution. However, the empirical distribution of stock returns is typically skewed, peaked around the mode, and has fat tails, implying that extreme events occur much more frequently than predicted by the normal distribution. Therefore, the traditional measures of market risk (e.g., variance or standard deviation) are not appropriate to approximate the maximum likely loss that a firm can expect under normal or highly volatile periods.[3]

Although the mean-variance criterion has been the basis for many academic papers and has had significant impact on the academic and nonacademic financial community, it is still subject to theoretical and empirical criticism. Arditti (1967), Arditti and Levy (1975), and Kraus and Litzenberger (1976) extend the standard portfolio theory to incorporate the effect of skewness on valuation. They present a three-moment model with unconditional skewness. Harvey and Siddique (2000) present an asset pricing model with conditional coskewness, where risk-averse investors prefer positively skewed assets to negatively skewed assets. Their results imply a preference for positive skewness: Investors should prefer stocks that are right-skewed to stocks that are left-skewed. Assets that decrease a portfolio's skewness (i.e., that make the portfolio returns more left-skewed) are less desirable and should command higher expected returns. Dittmar (2002) extends the three-moment asset pricing model using the restriction of decreasing absolute prudence (see Pratt and Zeckhauser (1987), Kimball (1993)). His findings suggest a preference for lower kurtosis. Investors are averse to kurtosis and prefer stocks with lower probability mass in the tails of the distribution to stocks with higher probability mass in the tails of the distribution. Assets that increase a portfolio's kurtosis (i.e., that make the portfolio returns more leptokurtic) are less desirable and should command higher expected returns. Since the magnitude of VaR becomes larger for negatively skewed and thicker-tailed asset distributions, the findings of the three- and four-moment asset pricing models indicate a positive relation between VaR and expected stock returns (i.e., the more a market index can potentially fall in value, the higher the expected return).

The paper is organized as follows. Section II presents alternative measures of market risk and describes our investigation of the risk-return tradeoff. Section III presents an economic framework that relates VaR to expected returns. Section IV presents the descriptive statistics of the data. Section V discusses the empirical results from time-series regressions. Section VI runs a battery of robustness checks. Section VII compares the relative performance of realized variance and VaR in

[3]Longin (2000), Neftci (2000), and Bali (2003) find that VaR provides good predictions of catastrophic market risks and performs surprisingly well in capturing both the rate of occurrence and the extent of extreme events in financial markets. However, the traditional measures of market risk such as conditional variance and standard deviation yield an inaccurate characterization of extreme movements in financial markets.

terms of their power to predict future market returns. Section VIII concludes the paper.

II. Measuring the Risk-Return Relationship

A. Alternative Risk Measures

1. Realized Variance

Following French, Schwert, and Stambaugh (1987), we calculate the variance of a market portfolio using various window sizes of return data:

$$\sigma_{k,t}^2 = \sum_{d=1}^{D_k} r_{k,d}^2 + 2\sum_{d=2}^{D_k} r_{k,d} \cdot r_{k,d-1}, \tag{1}$$

where $\sigma_{k,t}^2$ is the variance of index returns, D_k is the number of trading days over the past k months,[4] and $r_{k,d}$ is the portfolio's return on day d that resides within k months. The second term on the right-hand side adjusts for the autocorrelation in daily returns using the approach of French et al. (1987). Note that the realized variance measure given in equation (1) is not, strictly speaking, a variance measure since daily returns are not demeaned before taking the expectation. However, as pointed out by French et al. (1987), the impact of subtracting the means is trivial for short holding periods.

2. Nonparametric Value at Risk

VaR determines how much the value of a portfolio could decline over a given period of time with a given probability as a result of changes in market rates. For example, if the given period of time is one day and the given probability is 1%, the VaR measure would be an estimate of the decline in the portfolio value that could occur with a 1% probability over the next trading day. In other words, if the VaR measure is accurate, losses greater than the VaR measure should occur less than 1% of the time. In this paper, we use different confidence levels to check the robustness of VaR measures as an explanatory variable for the expected return on the market. The estimation is based on the lower tail of the actual empirical distribution. We use the past 1 to 6 months of daily returns to estimate alternative VaR measures from the empirical distribution. It should be noted that the original VaR measures are multiplied by -1 before running our regressions. The original maximum likely loss values are negative since they are obtained from the left tail of the distribution, but the downside risk measure, VaR_t, used in our regressions, is defined as -1 times the maximum likely loss. Therefore, the slope coefficients turn out to be positive, which gives the paper's central result that there is a positive and statistically significant relation between VaR and the excess return on the market.

[4]As in Ghysels et al. (2005), we use 1 to 6 months of past daily data to compute the rolling window variance estimates.

3. Parametric Value at Risk

There is substantial empirical evidence showing that the distribution of financial returns is typically skewed to the left, is peaked around the mean (leptokurtic), and has fat tails. The fat tails and negative skewness suggest that extreme outcomes happen much more frequently than would be predicted by the normal distribution, and the negative returns of a given magnitude have higher probabilities than positive returns of the same magnitude. This also suggests that the normality assumption can produce VaR numbers that are inappropriate measures of the true risk faced by individual firms. To account for skewness and excess kurtosis in the data, we use the skewed t (ST) distribution of Hansen (1994), which accounts for the nonnormality of returns and relatively infrequent events.

Hansen (1994) introduces a generalization of the Student t-distribution where asymmetries may occur, while maintaining the assumption of a zero mean and unit variance. The ST density that provides a flexible tool for modeling the empirical distribution of stock market returns exhibiting skewness and leptokurtosis is given by equation (2):

$$(2) \qquad f(z_t\,;\mu,\sigma,\upsilon,\lambda) = \begin{cases} bc\left(1+\dfrac{1}{\nu-2}\left(\dfrac{bz_t+a}{1-\lambda}\right)^2\right)^{-\frac{\upsilon+1}{2}} & \text{if } z_t < -\dfrac{a}{b} \\ bc\left(1+\dfrac{1}{\nu-2}\left(\dfrac{bz_t+a}{1+\lambda}\right)^2\right)^{-\frac{\upsilon+1}{2}} & \text{if } z_t \geq -\dfrac{a}{b} \end{cases},$$

where $z_t = (R_t - \mu)/\sigma$ is the standardized market return, and the constants a, b, and c are given by

$$(3) \qquad a = 4\lambda c\left(\frac{\nu-2}{\nu-1}\right), \qquad b^2 = 1+3\lambda^2-a^2,$$

$$c = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi(\nu-2)}\,\Gamma\left(\frac{\nu}{2}\right)}.$$

Hansen shows that this density is defined for $2 < \nu < \infty$ and $-1 < \lambda < 1$. This density has a single mode at $-a/b$, which is of opposite sign with the parameter λ. Thus, if $\lambda > 0$, the mode of the density is to the left of zero and the variable is skewed to the right, and vice versa when $\lambda < 0$. Furthermore, if $\lambda = 0$, Hansen's distribution reduces to the standardized t-distribution. If $\lambda = 0$ and $\nu = \infty$, it reduces to a normal density.[5]

---

[5]The parameters of the ST density are estimated by maximizing the log-likelihood function of R_t with respect to the parameters μ, σ, ν, and λ:

$$\begin{aligned}\log L = {} & n\ln b + n\ln\Gamma\left(\frac{\nu+1}{2}\right) - \frac{n}{2}\ln\pi - n\ln\Gamma(\nu-2) - n\ln\Gamma\left(\frac{\nu}{2}\right) \\ & - n\ln\sigma - \left(\frac{\nu+1}{2}\right)\sum_{t=1}^{n}\ln\left(1+\frac{d_t^2}{(\nu-2)}\right),\end{aligned}$$

A parametric approach to calculating VaR is based on the lower tail of the ST distribution. Specifically, we estimate the parameters of the ST density (μ, σ, ν, λ) using the past 1 to 6 months of daily data and then find the corresponding percentile of the estimated distribution. Assuming that $R_t \sim f_{\nu,\lambda}(z)$ follows a ST density, parametric VaR is the solution to

$$\int_{-\infty}^{\Gamma_{\mathrm{ST}}(\Phi)} f_{\nu,\lambda}(z)dz = \Phi, \tag{4}$$

where $\Gamma_{\mathrm{ST}}(\Phi)$ is the VaR threshold based on the ST density with a loss probability of Φ. Equation (4) indicates that VaR can be calculated by integrating the area under the probability density function of the ST distribution.

B. Time-Series Regressions

We investigate the intertemporal relation between downside risk and excess market return at the monthly frequency. The downside risk-return relationship we analyze in the paper takes the following form:

$$R_{t+1} = \alpha + \beta E_t(\mathrm{VaR}_{t+1}) + \gamma X_t + \varepsilon_{t+1}, \tag{5}$$

where R_{t+1} is the monthly excess return of the market portfolio, $E_t(\mathrm{VaR}_{t+1})$ is the conditional VaR of the market portfolio obtained from the daily index returns, and X_t is the vector of control variables that includes a set of macroeconomic variables proxying for business cycle fluctuations, the lagged excess return, and a dummy variable for the October 1987 crash.[6] We use various measures of the lagged realized VaR as a proxy for the expected conditional downside risk for the current period.[7] The slope coefficient β in equation (5) is expected to be positive and statistically significant.

We also test the usual form of the risk-return tradeoff by examining whether the relation between the conditional variance and the expected excess return is positive. We use the following discrete-time specification of Merton (1980):

$$R_{t+1} = \alpha + \beta E_t(\sigma^2_{t+1}) + \gamma X_t + \varepsilon_{t+1}, \tag{6}$$

where the coefficients α and β, according to Merton's ICAPM, should be 0 and equal to the relative risk aversion coefficient, respectively. Positive values of β imply the existence of a risk-return tradeoff, indicating that the expected returns are higher as the risk level for the market increases. Following Ghysels et al. (2005), we use various measures of the lagged realized variance as a proxy for $E_t(\sigma^2_{t+1})$.

where $d_t = (bz_t + a)/(1 - \lambda s)$ and s is a sign dummy taking the value of 1 if $bz_t + a < 0$ and $s = -1$ otherwise.

[6]To make sure that our results from estimating risk-return regressions are not due to model misspecification, we add to the regressions a set of control variables that have been used in the literature to capture the state variables that determine changes in the investment opportunity set.

[7]As discussed later in the paper, we conduct robustness checks where VaR measures that conditionally change over time are used in regressions. The results from the lagged realized VaR and from the conditional forecasts of VaR are found to be similar.

III. Economic Framework

The standard theory of portfolio choice determines the optimum asset mix by maximizing i) the expected risk premium per unit of risk in a mean-variance framework or ii) the expected value of a utility function approximated by the expected return and variance of the portfolio. In both cases, market risk of the portfolio is defined in terms of the variance (or standard deviation) of the portfolio's returns. Modeling portfolio risk with the traditional volatility measures implies that investors are concerned only about the average variation (and covariation) of individual stock returns, and they are not allowed to treat the negative and positive tails of the return distribution separately.

In what follows, we consider an investor who allocates her portfolio in order to maximize the expected utility of end-of-period wealth $U(W)$. We assume that the distribution of returns on the investor's portfolio of risky assets is nonsymmetrical and fat-tailed. The expected value of end-of-period wealth can be written as $\overline{W} = \sum_{i=1}^{n} q_i \overline{R}_i + q_f R_f$, where $\overline{R}_i$ is unity plus the expected rate of return on the ith risky asset, R_f is unity plus the rate of return on the riskless asset, q_i is the fraction of wealth allocated to the ith risky asset, and q_f is the fraction of wealth allocated to the riskless asset. We approximate the expected utility by a Taylor series expansion around the expected wealth. For this purpose, the utility function is expressed in terms of the wealth distribution, so that $\mathrm{E}\left[U(W)\right] = \int U(W) f(W) dW$, where $f(W)$ is the probability density function of the end-of-period wealth, which depends on the multivariate distribution of returns and on the vector of weights q. We now consider the infinite-order Taylor series expansion of the utility function

$$U(W) = \sum_{k=0}^{\infty} \frac{U^{(k)}(\overline{W})(W-\overline{W})^k}{k!}, \tag{7}$$

where $\overline{W} = \mathrm{E}(W)$ denotes the expected end-of-period wealth. Under rather mild conditions, the expected utility is given by

$$\mathrm{E}\left[U(W)\right] = \mathrm{E}\left[\sum_{k=0}^{\infty} \frac{U^{(k)}(\overline{W})(W-\overline{W})^k}{k!}\right] = \sum_{k=0}^{\infty} \frac{U^{(k)}(\overline{W})}{k!} \mathrm{E}\left[(W-\overline{W})^k\right]. \tag{8}$$

Therefore, the expected utility depends on all central moments of the distribution of the end-of-period wealth.

It should be noticed that the approximation of the expected utility by a Taylor series expansion is related to the investor's preference (or aversion) toward all moments of the distribution that are directly given by derivatives of the utility function. Scott and Horvath (1980) indicate that, under the assumptions of positive marginal utility, by decreasing absolute risk aversion at all wealth levels, together with strict consistency for moment preferences, one has $U^{(k)}(W) > 0$, $\forall W$ if k is odd and $U^{(k)}(W) < 0$, $\forall W$ if k is even. Further discussion on the conditions that yield such moment preferences or aversion can be found in Pratt and

Zeckhauser (1987), Kimball (1993), and Dittmar (2002). Focusing on terms up to the fourth one, we obtain

$$(9)\quad \mathrm{E}\left[U(W)\right] = U(\overline{W}) + U^{(1)}(\overline{W})\,\mathrm{E}\left[(W-\overline{W})\right] + \frac{1}{2}U^{(2)}(\overline{W})\,\mathrm{E}\left[(W-\overline{W})^2\right] + \frac{1}{3!}U^{(3)}(\overline{W})\,\mathrm{E}\left[(W-\overline{W})^3\right] + \frac{1}{4!}U^{(4)}(\overline{W})\,\mathrm{E}\left[(W-\overline{W})^4\right] + O(W^4),$$

where $O(W^4)$ is the Taylor remainder. We define the expected return, variance, skewness, and kurtosis of the end-of-period return, R_p, as

$$(10)\quad \mu_p = \mathrm{E}\left[R_p\right] = \overline{W},$$

$$(11)\quad \sigma_p^2 = \mathrm{E}\left[(R_p-\mu_p)^2\right] = \mathrm{E}\left[(W-\overline{W})^2\right],$$

$$(12)\quad s_p^3 = \mathrm{E}\left[(R_p-\mu_p)^3\right] = \mathrm{E}\left[(W-\overline{W})^3\right], \quad \text{and}$$

$$(13)\quad k_p^4 = \mathrm{E}\left[(R_p-\mu_p)^4\right] = \mathrm{E}\left[(W-\overline{W})^4\right].$$

Hence, the expected utility is simply approximated by the following preference function:

$$(14)\quad \mathrm{E}\left[U(W)\right] \approx U(\overline{W}) + \frac{1}{2}U^{(2)}(\overline{W})\sigma_p^2 + \frac{1}{3!}U^{(3)}(\overline{W})s_p^3 + \frac{1}{4!}U^{(4)}(\overline{W})k_p^4.$$

Under conditions established by Scott and Horvath (1980), the expected utility depends positively on expected returns and skewness and negatively on variance and kurtosis. Based on the constant absolute risk aversion (CARA) and constant relative risk aversion (CRRA) utility functions, we now show that an increase in VaR reduces the expected utility of wealth.

We first consider the CARA utility function, defined by: $U(W) = -\exp(-\theta W)$, where θ measures the investor's CARA. The approximation for the expected utility is given by

$$(15)\quad \mathrm{E}\left[U(W)\right] \approx -\exp(-\theta\overline{W})\left[1 + \frac{\theta^2}{2}\sigma_p^2 - \frac{\theta^3}{3!}s_p^3 + \frac{\theta^4}{4!}k_p^4\right].$$

Equation (15) indicates aversion to variance and kurtosis and preference for (positive) skewness since

$$\frac{\partial\,\mathrm{E}\left[U(W)\right]}{\partial\sigma_p^2} = -\exp(-\theta\overline{W})\frac{\theta^2}{2} < 0, \quad \frac{\partial\,\mathrm{E}\left[U(W)\right]}{\partial s_p^3} = \exp(-\theta\overline{W})\frac{\theta^3}{3!} > 0, \quad \text{and}$$

$$\frac{\partial\,\mathrm{E}\left[U(W)\right]}{\partial k_p^4} = -\exp(-\theta\overline{W})\frac{\theta^4}{4!} < 0.$$

These results imply aversion to VaR, $(\partial\,\mathrm{E}\left[U(W)\right])/(\partial \mathrm{VaR}_p) < 0$, since VaR for long positions (defined by the left tail of the return distribution) increases with variance and kurtosis and decreases with positive skewness (see Cornish and Fisher (1937)).

Similar results are obtained from the CRRA utility function given by: $U(W) = (W^{1-\theta})/(1-\theta)$, $(\theta > 0, \theta \neq 1)$. Since $\theta > 0$, the expected utility of wealth decreases with variance and kurtosis, whereas it increases with positive skewness, for example,

$$\frac{\partial \mathrm{E}[U(W)]}{\partial \sigma_p^2} = -\frac{\theta}{2}\overline{W}^{-(1+\theta)} < 0, \quad \frac{\partial \mathrm{E}[U(W)]}{\partial s_p^3} = \frac{\theta(1+\theta)}{3!}\overline{W}^{-(2+\theta)} > 0, \quad \text{and}$$

$$\frac{\partial \mathrm{E}[U(W)]}{\partial k_p^4} = -\frac{\theta(1+\theta)(2+\theta)}{4!}\overline{W}^{-(3+\theta)} < 0.$$

Since

$$\frac{\partial \mathrm{VaR}_p}{\partial \sigma_p^2} > 0, \quad \frac{\partial \mathrm{VaR}_p}{\partial s_p^3} < 0, \quad \text{and} \quad \frac{\partial \mathrm{VaR}_p}{\partial k_p^4} > 0$$

(see Cornish and Fisher (1937)), investors dislike VaR—for example, an increase in VaR reduces the expected utility of wealth, $(\partial \mathrm{E}[U(W)])/(\partial \mathrm{VaR}_p) < 0$. Hence, investors have an aversion to VaR that implies a positive relation between the VaR of a portfolio and the portfolio's expected return.

IV. Data

To capture the U.S. stock market returns, we use the value- and equal-weighted monthly returns on the NYSE/AMEX/Nasdaq index. As a robustness check, we also repeat our analysis for the NYSE/AMEX, NYSE, Nasdaq, and S&P 500 indices. We use returns from July 1962 to December 2005, except for the Nasdaq sample, which covers the period from January 1973 to December 2005. As a further robustness check, we also use a longer sample period of January 1926 to December 2005. In predictive regressions, the excess market return is defined as the difference between the index return and the risk-free rate. We use the 1-month Treasury bill return as the risk-free rate.

Panel A of Table 1 provides descriptive statistics for the value-weighted index returns. Panel A shows that the average monthly return is in the range of 0.97% for the NYSE/AMEX/Nasdaq and 1.05% for Nasdaq, which correspond to annualized returns of 11.64% and 12.60%, respectively. The unconditional standard deviations of monthly returns are in the range of 4.2% for the NYSE and 6.5% for the Nasdaq index. The skewness and kurtosis statistics are reported for testing the distributional assumption of normality. The skewness statistics for monthly returns are negative and significant at the 1% level. The kurtosis statistics are greater than 3.0 and statistically significant at the 1% level. Furthermore, the Jarque-Bera (1980) statistics strongly reject the distributional assumption of normality.[8]

[8]In Jarque-Bera (1980), $\mathrm{JB} = n[(S^2/6) + (K-3)^2/24]$, is a formal test statistic for testing whether returns are normally distributed, where n denotes the number of observations, S is skewness, and K is kurtosis. The test statistic, distributed as chi-square with two degrees of freedom, measures the difference of the skewness and kurtosis of the series from those of the normal distribution.

TABLE 1

Descriptive Statistics

Panel A of Table 1 shows summary statistics for the monthly return on the value-weighted NYSE/AMEX/Nasdaq, NYSE/AMEX, NYSE, Nasdaq, and S&P 500 index for the sample period from July 1962 to December 2005. Panel B shows summary statistics for Value at Risk (VaR_t), computed using rolling window estimation over various months (k). VaR_t is defined as -1 times the minimum NYSE/AMEX/Nasdaq index return observed during the last k months as of the end of each month t. Each month is assumed to have 21 trading days. We report the mean, median, standard deviation, maximum, minimum, skewness, kurtosis, and Jarque-Bera (1980) statistics. Panel C presents results from the AR(1) regressions, $\text{VaR}_{t+1} = \lambda + \rho\text{VaR}_t + \varepsilon_{t+1}$, where VaR_t is obtained from the past 1 to 6 months of daily data. For each regression in Panel C, we present the intercept (λ), AR(1) coefficient (ρ), Newey-West (1987) adjusted t-statistics (in parentheses), number of observations, and corrected R^2 values implied by the autocorrelation-heteroskedasticity adjusted t-statistics: $R^2 = t^2/(t^2 + (n-2))$, where t is the Newey-West (1987) adjusted t-statistic of the AR(1) coefficient and n is the number of observations.

Panel A. Monthly Index Returns

	NYSE/AMEX/Nasdaq	NYSE/AMEX	NYSE	Nasdaq	SP500
No. of obs.	522	522	522	396	522
Mean	0.970%	0.977%	0.980%	1.046%	0.974%
Median	1.250%	1.195%	1.195%	1.375%	1.085%
Std. dev.	4.405%	4.224%	4.198%	6.504%	4.273%
Maximum	16.56%	16.50%	16.81%	21.98%	16.81%
Minimum	−22.53%	−21.81%	−21.62%	−27.11%	−21.58%
Skewness	−0.4547	−0.3743	−0.3495	−0.4792	−0.3172
Kurtosis	5.0203	5.2019	5.1738	4.7683	4.9166
Jarque-Bera	106.76	117.64	113.41	66.75	88.65

Panel B. Value at Risk over Various Horizons

k	1	2	3	4	5	6
No. of obs.	521	520	519	518	517	516
Mean	0.015	0.019	0.021	0.023	0.024	0.026
Median	0.013	0.016	0.017	0.018	0.020	0.022
Std. dev.	0.011	0.014	0.016	0.017	0.019	0.020
Maximum	0.171	0.171	0.171	0.171	0.171	0.171
Minimum	0.002	0.004	0.005	0.005	0.008	0.008
Skewness	6.240	5.977	5.560	5.252	4.934	4.699
Kurtosis	77.405	60.826	49.431	42.408	36.584	32.650

Panel C. AR(1) Regressions of VaR

k	1	2	3	4	5	6
Intercept	0.011	0.007	0.005	0.004	0.004	0.003
	(9.03)	(6.08)	(4.85)	(4.06)	(3.41)	(2.87)
AR(1) coeff.	0.286	0.610	0.754	0.805	0.843	0.871
	(4.99)	(10.15)	(14.07)	(14.95)	(15.51)	(15.99)
No. of obs.	521	520	519	518	517	516
Corrected R^2	4.6%	16.6%	27.7%	30.2%	31.8%	33.2%

Panel B of Table 1 shows summary statistics for VaR computed using a rolling window estimation over various months denoted by k. VaR is defined as -1 times the minimum NYSE/AMEX/Nasdaq index return observed during the last k months of daily data as of the end of each month t. Furthermore, each month is assumed to have 21 trading days. For example, VaR for the past 4 months is computed as the lowest return observed during the last 84 days. Observe that the distributions of various VaRs are skewed to the right and have fatter tails than the normal distribution.

Panel C of Table 1 presents results from the AR(1) regressions for alternative measures of VaR, $\text{VaR}_{t+1} = \lambda + \rho\text{VaR}_t + \varepsilon_{t+1}$, where VaR_t is obtained from the past 1 to 6 months of daily data. For each regression in Panel C, we present the intercept (λ), AR(1) coefficient (ρ), their Newey-West (1987) adjusted t-statistics in parentheses, the number of observations, and the corrected R^2 values implied by the autocorrelation-heteroskedasticity adjusted t-statistics: $R^2 = t^2/(t^2 + (n-2))$, where t is the Newey-West (1987) adjusted t-statistic of the AR(1) coefficient, and n is the number of observations. As shown in Panel C, the AR(1) coefficients are in the range of 0.29 to 0.87, and they are all statistically significant at the 1% level. The t-statistics and the corrected R^2 values indicate that the realized VaR measures are highly persistent and motivate the use of VaR_t as a proxy for $E_t(\text{VaR}_{t+1})$.

A series of papers argue that the stock market can be predicted by macroeconomic variables associated with business cycle fluctuations. The commonly chosen variables include default spread (DEF_t), term spread (TERM_t), dividend price ratio (DP_t), and the detrended riskless rate (RREL_t).[9] We define DEF_t as the change in the difference between the yields on BAA- and AAA-rated corporate bonds, and TERM_t as the change in the difference between the yields on the 10-year Treasury bond and the 1-month Treasury bill. RREL_t is defined as the difference between the 1-month T-bill rate and its 12-month backward moving average.[10] To be consistent with earlier studies on ICAPM, we use the aforementioned macroeconomic variables in our risk-return regressions and investigate how incorporating these variables into the predictive regressions affects the intertemporal relation between downside risk and expected stock returns.[11]

V. Empirical Results

A. Results from Nonparametric Value at Risk

Table 2 presents the first set of empirical results from the time-series regressions of the value-weighted excess market return on nonparametric VaR (VaR_t) computed using daily returns observed over the past 1 to 6 months. The first column in each panel shows the number of months used to compute VaR. We assume that each month has 21 trading days. Therefore, at the 1-month horizon, VaR is defined as the minimum daily return observed during the past 21 days; hence, it corresponds to 4.76% VaR. At the 2-month horizon, VaR is defined as the minimum daily return observed during the past 42 days; hence, it can be viewed as 2.38% VaR. Similarly, at the 5-month horizon, VaR is defined as the minimum

[9]See, e.g., Campbell (1987), Fama and French (1989), Ferson and Harvey (1991), Tallarini and Zhang (2005), and Bali (2008), who test the predictive power of these variables for expected stock returns.

[10]The time-series data on monthly 10-year Treasury bond yields and on BAA- and AAA-rated corporate bond yields are available from the Federal Reserve statistics release Web site. We obtain the dividend price ratio using the CRSP value-weighted index return with and without dividends based on the formula given in Fama and French (1988). Finally, we obtain the 1-month Treasury bill rate from Kenneth French's online data library (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). At an earlier stage of the study, we proxy the short-term interest rate with the 3-month Treasury yields reported on the Federal Reserve Web site (http://www.federalreserve.gov/releases/h15/data.htm). Note that our results are not sensitive to the choice of 1- versus 3-month T-bill rates.

[11]At an earlier stage of the study, we use the default spread (DEF) and term spread (TERM) as the difference between two interest rate levels. When DEF and TERM are defined in the level of interest rates (instead of the change in interest rates), the slope coefficients on VaR turn out to be very similar to those reported in our tables. However, the coefficients on DEF and TERM are found to be statistically insignificant. This provides further evidence that innovations in macroeconomic variables (instead of the levels) generate a better proxy for state variables capturing shifts in the investment opportunity set.

daily return observed during the past 105 days; hence, it is 0.95% VaR, etc. In predictive regressions, the dependent variable is the 1-month-ahead value-weighted excess return on the NYSE/AMEX/Nasdaq index, R_{t+1}. The independent variables are VaR, a dummy variable that takes the value of 1 in October 1987, and 0 otherwise, the lagged excess market return, R_t, and the macroeconomic variables. For each parameter estimate, we present the Newey-West (1987) adjusted t-statistic in parentheses.[12]

TABLE 2

Value-Weighted Market Return and Nonparametric Value at Risk

Table 2 presents results from the time-series regressions of the value-weighted excess market return on nonparametric Value at Risk (VaR), defined as the minimum daily index return observed during the past 1 to 6 months. The original VaRs are multiplied by −1 before running our regressions. Value-weighted excess return, R_t, is defined as the return on the value-weighted NYSE/AMEX/Nasdaq index minus the 1-month Treasury bill rate. A dummy takes the value of 1 in October 1987, and 0 otherwise. DEF_t is the default spread calculated as the change in the difference between the yields on BAA- and AAA-rated corporate bonds. $TERM_t$ is the term spread calculated as the change in the difference between the yields on the 10-year Treasury bond and the 1-month Treasury bill. $RREL_t$ is the stochastically detrended riskless rate defined as the 1-month Treasury bill rate minus its 12-month backward-moving average. DP_t is the aggregate dividend yield. In each regression, the dependent variable is the 1-month-ahead excess market return, R_{t+1}. The first row gives the estimated coefficients. The second row gives the Newey-West (1987) adjusted t-statistics in parentheses. The R^2 values are reported in the last column.

Panel A. Mean-VaR Tradeoff

k	Constant	VaR_t	R_t	Dummy	R^2
1	−0.003	0.498	0.078	−0.141	1.6%
	(−0.72)	(2.25)	(1.71)	(−4.51)	
2	−0.004	0.491	0.061	−0.142	2.5%
	(−1.48)	(4.23)	(1.45)	(−7.77)	
3	−0.003	0.391	0.044	−0.129	2.4%
	(−1.07)	(3.72)	(1.11)	(−7.45)	
4	−0.003	0.368	0.040	−0.126	2.6%
	(−1.17)	(4.19)	(1.00)	(−7.96)	
5	−0.002	0.266	0.036	−0.111	1.9%
	(−0.57)	(3.02)	(0.90)	(−7.01)	
6	−0.001	0.221	0.037	−0.104	1.7%
	(−0.32)	(2.61)	(0.91)	(−6.69)	

Panel B. Mean-VaR Tradeoff with Macro Variables

k	Constant	VaR_t	R_t	Dummy	$RREL_t$	$TERM_t$	DEF_t	DP_t	R^2
1	−0.012	0.451	0.063	−0.141	−0.467	−0.739	3.381	0.346	4.9%
	(−1.58)	(2.01)	(1.34)	(−4.44)	(−2.57)	(−2.44)	(2.00)	(1.63)	
2	−0.014	0.468	0.051	−0.146	−0.461	−0.761	2.977	0.357	5.7%
	(−1.98)	(3.84)	(1.17)	(−7.63)	(−2.57)	(−2.53)	(1.76)	(1.67)	
3	−0.013	0.373	0.035	−0.134	−0.460	−0.751	3.159	0.349	5.6%
	(−1.78)	(3.48)	(0.84)	(−7.38)	(−2.55)	(−2.52)	(1.89)	(1.62)	
4	−0.013	0.351	0.031	−0.132	−0.443	−0.759	3.034	0.352	5.7%
	(−1.80)	(3.70)	(0.76)	(−7.52)	(−2.44)	(−2.56)	(1.79)	(1.63)	
5	−0.011	0.238	0.025	−0.114	−0.457	−0.711	3.408	0.326	5.0%
	(−1.42)	(2.42)	(0.62)	(−6.32)	(−2.51)	(−2.38)	(2.02)	(1.51)	
6	−0.010	0.200	0.024	−0.108	−0.463	−0.713	3.532	0.323	4.9%
	(−1.33)	(2.16)	(0.62)	(−6.23)	(−2.54)	(−2.38)	(2.11)	(1.50)	

As shown in Panel A of Table 2, after controlling for the lagged market return and the crash dummy, the slope coefficients on VaR_t are positive and highly significant. Specifically, the Newey-West (1987) t-statistics of the VaR slopes are in the range of 2.25 to 4.23. The dummy variable controls for the October 1987 crash.

[12]We use six lags when computing the Newey-West (1987) standard errors.

The coefficient on the dummy variable is negative and significant, indicating that there would be a misspecification error if we had not used it as a right-hand side variable.[13]

Our main finding can be interpreted by considering the changes in expected excess returns as a result of a one-standard-deviation change in the average VaR measure. For example, for the 1-month horizon ($k = 1$), moving from average VaR minus one standard deviation (i.e., $1.5\% - 1.1\% = 0.4\%$) to average VaR plus one standard deviation (i.e., $1.5\% + 1.1\% = 2.6\%$) increases expected excess returns by 1.10% per month ($2.2\% \times 0.498 = 1.10\%$).

We further investigate the relation between downside risk and expected returns after controlling for macroeconomic variables known to forecast the stock market. As shown by Merton (1973), the hedging demand component (γX_t) in equations (5) and (6) captures the investor's motive to hedge against unfavorable shifts in the investment opportunity set. Thus, we include macroeconomic variables that have been shown in the literature to capture state variables that determine the investment opportunity set.

Panel B of Table 2 shows that even after controlling for the macroeconomic variables, VaR_t has a positive and significant coefficient, indicating that there is a robust and significantly positive relationship between downside risk and expected returns. The strong positive relation holds for alternative measures of VaR obtained from the past 1 to 6 months of daily data. Consistent with the earlier research (e.g., Ghysels et al. (2005), Bali, Cakici, Yan, and Zhang (2005)), the R^2 values are small, in the range of 4.9% to 5.7%.

B. Results from Parametric Value at Risk

Table 3 presents the parameter estimates from the monthly regressions of the value-weighted Center for Research in Security Prices (CRSP) index returns on the lagged VaR, which is calculated parametrically based on the lower tail of the ST distribution, VaR_t^p. The magnitude and statistical significance of the slope coefficients turn out to be very similar to our findings in Table 2. As shown in Table 3, VaR_t^p has positive and highly significant coefficients. Specifically, the Newey-West (1987) t-statistics of the slope coefficients on VaR_t^p range from 1.96 to 3.70. This result is robust across different measures of VaR computed with the ST density using the past 1 to 6 months of daily data. Overall, the results from alternative measures of nonparametric and parametric VaR turn out to be similar, indicating a significantly positive relation between downside risk and expected return on the market.

C. Results from the Equal-Weighted Market Portfolio

We have so far examined the significance of downside risk based on the value-weighted portfolio of NYSE/AMEX/Nasdaq stocks. We will now repeat our analyses for the equal-weighted market index. Stocks with smaller (bigger)

---

[13]We also repeat our analysis by eliminating the month of October 1987. Since the qualitative results turn out to be very similar to those reported in our tables, we do not present them here. They are available from the authors.

TABLE 3

Value-Weighted Market Return and Parametric Value at Risk

Table 3 presents results from the time-series regressions of the value-weighted excess market return on parametric Value at Risk (VaR_t^P), defined based on the lower tail of the skewed t density using daily returns over the past 1 to 6 months. The original VaRs are multiplied by -1 before running our regressions. Control variables are defined in Table 2. In each regression, the dependent variable is the 1-month-ahead excess market return, R_{t+1}. The first row gives the estimated coefficients. The second row gives the Newey-West (1987) adjusted t-statistics in parentheses. The R^2 values are reported in the last column.

k	Constant	VaR_t^P	R_t	Dummy	$RREL_t$	$TERM_t$	DEF_t	DP_t	R^2
1	−0.012	0.399	0.061	−0.139	−0.472	−0.744	3.395	0.347	4.8%
	(−1.59)	(1.96)	(1.31)	(−4.39)	(−2.60)	(−2.46)	(1.98)	(1.64)	
2	−0.014	0.448	0.049	−0.142	−0.476	−0.773	3.029	0.356	5.4%
	(−1.93)	(3.63)	(1.13)	(−7.27)	(−2.64)	(−2.54)	(1.77)	(1.67)	
3	−0.013	0.374	0.036	−0.131	−0.476	−0.761	3.180	0.345	5.3%
	(−1.72)	(3.43)	(0.86)	(−7.39)	(−2.63)	(−2.54)	(1.90)	(1.61)	
4	−0.013	0.382	0.033	−0.131	−0.462	−0.774	3.055	0.349	5.6%
	(−1.80)	(3.70)	(0.80)	(−7.40)	(−2.54)	(−2.59)	(1.81)	(1.62)	
5	−0.010	0.252	0.026	−0.113	−0.474	−0.722	3.450	0.317	4.9%
	(−1.36)	(2.31)	(0.64)	(−6.11)	(−2.61)	(−2.40)	(2.04)	(1.48)	
6	−0.009	0.214	0.026	−0.108	−0.479	−0.722	3.563	0.311	4.8%
	(−1.26)	(2.07)	(0.62)	(−6.03)	(−2.64)	(−2.40)	(2.13)	(1.45)	

market capitalization are weighted more heavily in the equal-weighted (value-weighted) index. We expect VaR to be even more powerful in forecasting the future equal-weighted returns because the return distributions of small stocks exhibit higher peaks, fatter tails, and more outliers on the left or right tail than do the distributions of bigger stocks.

Table 4 presents the parameter estimates and the Newey-West (1987) t-statistics for alternative measures of nonparametric VaR that are used to predict the excess return on the CRSP equal-weighted index. The slope coefficients on VaR are positive and highly significant, with the t-statistics falling in the range of 2.83 to 4.77 even when macroeconomic variables are used as control variables. These results also provide supporting evidence for a stronger relation between VaR and expected return on the equal-weighted index.

TABLE 4

Equal-Weighted Market Return and Nonparametric Value at Risk

Table 4 presents results from the time-series regressions of the equal-weighted excess market return on nonparametric Value at Risk (VaR_t), defined as the minimum daily index return observed during the past 1 to 6 months. The original VaRs are multiplied by -1 before running our regressions. Control variables are defined in Table 2. In each regression, the dependent variable is the 1-month-ahead excess market return, R_{t+1}. The first row gives the estimated coefficients. The second row gives the Newey-West (1987) adjusted t-statistics in parentheses. The R^2 values are reported in the last column.

k	Constant	VaR_t	R_t	Dummy	$RREL_t$	$TERM_t$	DEF_t	DP_t	R^2
1	−0.011	1.064	0.298	−0.081	−0.745	−1.293	3.184	0.282	10.4%
	(−1.18)	(2.83)	(6.07)	(−2.53)	(−3.74)	(−3.49)	(1.31)	(1.20)	
2	−0.008	0.722	0.242	−0.062	−0.797	−1.326	2.732	0.249	10.2%
	(−0.90)	(3.59)	(5.47)	(−2.81)	(−3.96)	(−3.52)	(1.12)	(1.04)	
3	−0.010	0.761	0.233	−0.066	−0.789	−1.336	2.535	0.256	10.9%
	(−1.16)	(4.29)	(5.45)	(−3.14)	(−3.89)	(−3.63)	(1.06)	(1.05)	
4	−0.011	0.709	0.222	−0.065	−0.768	−1.380	2.307	0.259	10.9%
	(−1.17)	(4.77)	(5.21)	(−3.22)	(−3.74)	(−3.76)	(0.97)	(1.04)	
5	−0.005	0.445	0.215	−0.041	−0.753	−1.278	3.356	0.225	9.9%
	(−0.58)	(3.33)	(5.14)	(−2.29)	(−3.65)	(−3.51)	(1.38)	(0.92)	
6	−0.005	0.387	0.218	−0.035	−0.744	−1.244	3.554	0.231	9.8%
	(−0.52)	(2.91)	(5.27)	(−1.96)	(−3.57)	(−3.39)	(1.45)	(0.94)	

D. Conditional Value at Risk

Before we generate the conditional VaR measure, we investigate the information content of VaR. As discussed in Section II, VaR is a nonlinear function of volatility, skewness, and kurtosis. Table 5 shows that past skewness has almost no power to predict future skewness. The Newey-West (1987) t-statistics of the slope coefficients on past skewness are in the range of 0.78 to 1.76 (see Chen, Hong, and Stein (2001) for a detailed analysis of the forecastability of skewness). On the other hand, the slope coefficients on past VaR have the correct sign for all horizons, and the Newey-West t-statistics are in the range of −2.40 to −5.49 (except for the 1-month horizon).[14] However, we find that (although not presented in the paper to save space) VaR has a little forecasting power for future kurtosis.

TABLE 5

Forecasting Skewness

Table 5 presents the parameter estimates, the Newey-West (1987) t-statistics in parentheses, and the R^2 values from the regression of 1-month-ahead skewness ($SKEW_{t+1}$) on past skewness ($SKEW_t$), past VaR (VaR_t), past return (R_t), and past volatility (σ_t). The original VaRs are multiplied by −1 before running our regressions.

k	Constant	$SKEW_t$	VaR_t	R_t	σ_t	R^2
1	−0.115	0.060	−8.339	−2.305	29.323	4.5%
	(−2.05)	(0.78)	(−1.49)	(−3.68)	(−2.06)	
2	−0.098	0.092	−6.315	−1.981	24.039	5.1%
	(−1.80)	(1.64)	(−2.80)	(−3.73)	(3.25)	
3	−0.089	0.082	−8.661	−1.778	29.714	6.7%
	(−1.68)	(1.47)	(−5.49)	(−3.38)	(4.53)	
4	−0.092	0.092	−5.088	−1.735	21.469	5.3%
	(−1.66)	(1.60)	(−3.54)	(−3.17)	(3.58)	
5	−0.088	0.101	−4.037	−1.718	18.784	5.2%
	(−1.57)	(1.76)	(−2.40)	(−3.07)	(3.19)	
6	−0.080	0.097	−4.888	−1.689	20.514	5.7%
	(−1.43)	(1.66)	(−3.09)	(−3.02)	(3.45)	

As mentioned earlier, we approximate the conditional VaR by the lagged realized VaR (i.e., $E_t(\mathrm{VaR}_{t+1}) \equiv \mathrm{VaR}_t$), to test the intertemporal relation between downside risk and excess market return as shown in equation (5). This approximation is justified by the fact that the VaR is highly persistent. As discussed in Section IV, we present the parameter estimates, the Newey-West (1987) adjusted t-statistics, and the R^2 values implied by the autocorrelation-heteroskedasticity adjusted t-statistics from an AR(1) specification of the realized VaR, $\mathrm{VaR}_{t+1} = \lambda + \rho \mathrm{VaR}_t + \varepsilon_{t+1}$. As shown in Panel C of Table 1, the AR(1) coefficients are in the range of 0.29 to 0.87, and they are all statistically significant at the 1% level.

Since the actual realizations of VaR measures are known to be conditionally changing over time, in this section we use time-varying conditional measures of downside risk. In order to come up with a more accurate conditional expectation of VaR, we consider a more general AR(p) specification:

$$\mathrm{VaR}_{t+1} = \lambda + \sum_{i=0}^{p-1} \rho_i \mathrm{VaR}_{t-i} + \varepsilon_{t+1}, \tag{16}$$

[14]In regressions, VaR is used in absolute value terms, and hence we expect a negative relation between VaR and skewness.

where p denotes the order of autoregressive process. For each measure of VaR, the optimal lag length p is determined based on the Akaike information criterion (AIC) and Schwarz Bayesian criterion (SBC). For 21-day VaR, the optimal lag length is found to be four, and alternative measures of VaR have an optimal lag length in the range of three to five. Hence, we use $p = 4$ when we compute the conditional measures. We should also note that the results are robust for all lag lengths between $p = 2$ and $p = 5$. Therefore, assuming that investors' information set as of time t consists of the lagged VaRs, we estimate the conditional VaR as the explained portion of the regression shown in equation (16), for example, $E_t(\text{VaR}_{t+1}) = \widehat{\lambda} + \sum_{i=0}^{p-1} \widehat{\rho}_i \text{VaR}_{t-i}$.

Table 6 presents results from the time-series regressions of the 1-month-ahead excess market return, R_{t+1}, on the conditional VaR, $E_t(\text{VaR}_{t+1})$, a dummy variable that takes the value of 1 in October 1987, and 0 otherwise, lagged excess market return, R_t, and macroeconomic variables. Here, $E_t(\text{VaR}_{t+1})$ has positive and highly significant coefficients. Specifically, the Newey-West (1987) t-statistics of the slope coefficients on $E_t(\text{VaR}_{t+1})$ are in the range of 2.28 to 3.89. The results in Table 6 indicate that there exists a positive and significant link between expected return and conditional VaR, $E_t(\text{VaR}_{t+1})$, for all estimation windows from 1 to 6 months. Furthermore, the results are similar to those reported in our earlier tables using the lagged realized VaR as a proxy for the conditional measure of downside risk. Thus, we conclude that the lagged VaR is a good proxy for the conditional expectation of future VaR.

TABLE 6

Excess Market Return and Conditional Value at Risk

Table 6 presents results from the time-series regressions of the excess market return on the conditional VaR, $E_t(\text{VaR}_{t+1})$, estimated using equation (16). The original conditional VaRs are multiplied by −1 before running our regressions. In each regression, the dependent variable is the 1-month-ahead excess market return, R_{t+1}. The first row gives the estimated coefficients. The second row gives the Newey-West (1987) adjusted t-statistics in parentheses. The R^2 values are reported in the last column.

k	Constant	$E_t(\text{VaR}_{t+1})$	R_t	Dummy	RREL_t	TERM_t	DEF_t	DP_t	R^2
1	−0.022	1.098	0.040	−0.113	−0.464	−0.746	3.222	0.353	5.1%
	(−2.25)	(2.53)	(0.96)	(−6.70)	(−2.55)	(−2.50)	(1.90)	(1.63)	
2	−0.019	0.733	0.058	−0.165	−0.450	−0.736	2.996	0.354	5.6%
	(−2.33)	(3.27)	(1.31)	(−6.07)	(−2.52)	(−2.46)	(1.77)	(1.67)	
3	−0.015	0.476	0.042	−0.141	−0.465	−0.738	3.195	0.340	5.5%
	(−1.94)	(3.44)	(0.96)	(−7.24)	(−2.59)	(−2.46)	(1.92)	(1.58)	
4	−0.015	0.453	0.031	−0.136	−0.448	−0.746	3.069	0.333	5.9%
	(−2.01)	(3.89)	(0.76)	(−7.63)	(−2.48)	(−2.52)	(1.81)	(1.55)	
5	−0.012	0.285	0.026	−0.116	−0.456	−0.708	3.430	0.320	5.1%
	(−1.54)	(2.51)	(0.63)	(−6.37)	(−2.50)	(−2.37)	(2.03)	(1.50)	
6	−0.011	0.239	0.028	−0.110	−0.460	−0.712	3.542	0.319	5.0%
	(−1.45)	(2.28)	(0.68)	(−6.25)	(−2.53)	(−2.37)	(2.12)	(1.49)	

VI. Robustness Checks

A. Different Indices

The positive relation between VaR_t and expected returns on the NYSE/AMEX/Nasdaq index may well be due to smaller, illiquid, and lower-priced stocks trading in certain exchanges. Therefore, in Table 7 we examine the relation between downside risk and expected returns for other stock market indices.

TABLE 7

Various Indices

Table 7 presents results from the time-series regressions of the excess market return on nonparametric Value at Risk (VaR_t). The original VaRs are multiplied by −1 before running our regressions. The dependent variable is defined as the return on the NYSE/AMEX (Panel A), NYSE (Panel B), Nasdaq (Panel C), and S&P 500 (Panel D) index minus the 1-month Treasury bill rate. In each regression, the first row gives the estimated coefficients. The second row gives the Newey-West (1987) adjusted t-statistics in parentheses. The R^2 values are reported in the last column.

k	Constant	VaR_t	R_t	Dummy	$RREL_t$	$TERM_t$	DEF_t	DP_t	R^2
Panel A. NYSE/AMEX									
1	−0.009	0.393	0.037	−0.146	−0.452	−0.746	3.500	0.274	4.9%
	(−1.31)	(1.88)	(0.79)	(−4.51)	(−2.57)	(−2.51)	(2.12)	(1.42)	
2	−0.011	0.412	0.027	−0.150	−0.444	−0.764	3.178	0.280	5.7%
	(−1.68)	(3.82)	(0.60)	(−8.15)	(−2.56)	(−2.59)	(1.93)	(1.46)	
3	−0.010	0.343	0.015	−0.141	−0.439	−0.752	3.330	0.279	5.7%
	(−1.56)	(3.57)	(0.34)	(−8.04)	(−2.52)	(−2.58)	(2.05)	(1.44)	
4	−0.010	0.309	0.011	−0.136	−0.426	−0.760	3.237	0.282	5.7%
	(−1.56)	(3.63)	(0.26)	(−8.08)	(−2.42)	(−2.62)	(1.97)	(1.46)	
5	−0.008	0.211	0.007	−0.120	−0.439	−0.715	3.550	0.267	5.1%
	(−1.21)	(2.26)	(0.16)	(−6.62)	(−2.49)	(−2.44)	(2.16)	(1.38)	
6	−0.007	0.180	0.006	−0.115	−0.444	−0.718	3.654	0.266	5.0%
	(−1.14)	(2.06)	(0.14)	(−6.62)	(−2.52)	(−2.45)	(2.24)	(1.37)	
Panel B. NYSE									
1	−0.009	0.397	0.030	−0.149	−0.456	−0.745	3.408	0.272	4.9%
	(−1.31)	(1.91)	(0.65)	(−4.57)	(−2.61)	(−2.53)	(2.11)	(1.43)	
2	−0.010	0.404	0.020	−0.152	−0.448	−0.764	3.111	0.278	5.7%
	(−1.65)	(3.77)	(0.65)	(−8.15)	(−2.61)	(−2.61)	(1.93)	(1.46)	
3	−0.010	0.335	0.008	−0.142	−0.444	−0.751	3.259	0.278	5.7%
	(−1.53)	(3.54)	(0.18)	(−8.11)	(−2.57)	(−2.60)	(2.05)	(1.45)	
4	−0.010	0.301	0.004	−0.137	−0.431	−0.758	3.172	0.280	5.7%
	(−1.53)	(3.60)	(0.10)	(−8.16)	(−2.47)	(−2.64)	(1.97)	(1.46)	
5	−0.008	0.206	0.000	−0.121	−0.444	−0.714	3.474	0.265	5.1%
	(−1.18)	(2.21)	(0.00)	(−6.65)	(−2.54)	(−2.46)	(2.16)	(1.38)	
6	−0.007	0.176	−0.001	−0.116	−0.450	−0.718	3.576	0.264	5.0%
	(−1.10)	(2.02)	(−0.02)	(−6.66)	(−2.58)	(−2.47)	(2.24)	(1.37)	
Panel C. Nasdaq									
1	−0.011	1.247	0.341	−0.062	−0.813	−1.309	2.372	0.262	12.4%
	(−0.78)	(2.41)	(5.24)	(−1.63)	(−3.53)	(−3.63)	(0.92)	(0.88)	
2	−0.003	0.743	0.270	−0.039	−0.853	−1.344	1.780	0.138	11.5%
	(−0.22)	(2.71)	(4.17)	(−1.61)	(−3.62)	(−3.67)	(0.68)	(0.47)	
3	−0.003	0.689	0.258	−0.038	−0.847	−1.375	1.799	0.130	11.5%
	(−0.27)	(2.71)	(4.16)	(−1.60)	(−3.58)	(−3.83)	(0.71)	(0.45)	
4	−0.005	0.700	0.252	−0.040	−0.808	−1.419	1.538	0.137	11.8%
	(−0.40)	(3.50)	(4.19)	(−1.81)	(−3.41)	(−3.97)	(0.61)	(0.46)	
5	0.003	0.425	0.235	−0.024	−0.770	−1.331	2.614	0.050	10.4%
	(0.26)	(2.38)	(3.99)	(−1.18)	(−3.23)	(−3.79)	(1.01)	(0.17)	
6	0.006	0.323	0.232	−0.017	−0.756	−1.292	2.830	0.026	9.8%
	(0.47)	(1.76)	(3.94)	(−0.87)	(−3.14)	(−3.66)	(1.08)	(0.09)	
Panel D. S&P 500									
1	−0.010	0.417	0.004	−0.169	−0.457	−0.717	3.438	0.291	5.0%
	(−1.35)	(2.12)	(0.08)	(−5.16)	(−2.52)	(−2.40)	(2.17)	(1.39)	
2	−0.011	0.396	−0.008	−0.167	−0.450	−0.735	3.135	0.298	5.7%
	(−1.62)	(3.69)	(−0.19)	(−8.48)	(−2.52)	(−2.49)	(1.98)	(1.43)	
3	−0.010	0.314	−0.021	−0.155	−0.450	−0.726	3.305	0.293	5.6%
	(−1.45)	(3.35)	(−0.49)	(−8.52)	(−2.51)	(−2.48)	(2.11)	(1.39)	
4	−0.010	0.294	−0.024	−0.152	−0.436	−0.731	3.216	0.296	5.7%
	(−1.47)	(3.37)	(−0.55)	(−8.44)	(−2.42)	(−2.51)	(2.03)	(1.40)	
5	−0.008	0.202	−0.029	−0.135	−0.451	−0.689	3.513	0.273	5.1%
	(−1.12)	(2.10)	(−0.67)	(−6.90)	(−2.49)	(−2.35)	(2.23)	(1.30)	
6	−0.008	0.175	−0.030	−0.131	−0.456	−0.694	3.606	0.271	5.1%
	(−1.06)	(1.93)	(−0.69)	(−6.97)	(−2.52)	(−2.36)	(2.30)	(1.29)	

Table 7 reports results for the value-weighted NYSE/AMEX, NYSE, Nasdaq, and S&P 500 index portfolios. For all market indices, there is a positive and highly significant relation between VaR_t and excess market return. As shown in Table 7, after controlling for the lagged market return, the October 1987 crash, and macroeconomic variables, the slope coefficients on VaR have Newey-West (1987) t-statistics ranging from 1.88 to 3.82 for NYSE/AMEX (Panel A), from 1.91 to 3.77 for NYSE (Panel B), from 1.76 to 3.50 for Nasdaq (Panel C), and from 1.93 to 3.69 for S&P 500 (Panel D). Although not presented in the paper to save space, the t-statistics are even larger when we exclude the macroeconomic variables. Thus, we conclude that the strong positive relation between VaR and expected return is robust across different stock market indices.

B. Alternative Measures of Downside Risk

VaR provides information about the left tail of the empirical return distribution; however, it is not the only measure of downside risk. If downside risk is an important determinant of expected returns, we expect other proxies of downside risk to perform well in predictive regressions too. In this section, we conduct an empirical analysis of the various left-tail risk measures.

An important example for a risk measure of this kind is "expected shortfall" originally proposed by Artzner, Delbaen, Eber, and Heath (1999). ES is defined as the conditional expectation of loss given that the loss is beyond the VaR level. That is, when the distributions of losses are continuous, ES at the $100(1-\Phi)\%$ confidence level is defined by

$$\text{ES}_{\Phi}(R_t) \;=\; \text{E}\left[R_t | R_t \leq \text{VaR}_{\Phi}(R_t)\right]. \tag{17}$$

Equation (17) can be viewed as a mathematical transcription of the concept "average loss in the worst $100\Phi\%$ cases."

In addition to ES that measures the mean of losses larger than VaR, we also compute the variance of losses larger than VaR and call it TR:

$$\text{TR}_{\Phi}(R_t) \;=\; \text{E}\left[\left(R_t - \text{E}\left(R_t | R_t \leq \text{VaR}_{\Phi}(R_t)\right)\right)^2 | R_t \leq \text{VaR}_{\Phi}(R_t)\right]. \tag{18}$$

We consider the 2.5% and 5% tail risk ($\text{TR}_t^{2.5\%}$ and $\text{TR}_t^{5\%}$) and the 2.5% and 5% expected shortfall ($\text{ES}_t^{2.5\%}$ and $\text{ES}_t^{5\%}$) as alternative proxies for downside risk. In our empirical analysis, we define the 2.5% (5%) TR as the sum of squared deviations of the lowest 2.5 percentile (5 percentile) of the NYSE/AMEX/Nasdaq index returns from the mean of index returns during the last 100 days. Similarly, we define the 2.5% (5%) ES as the average of the lowest 2.5 percentile (5 percentile) of the NYSE/AMEX/Nasdaq index returns observed during the last 100 days as of the end of month t.[15,16]

[15] When computing ES and TR, we need a large number of observations so that we can find the mean and variance of extreme observations beyond some VaR threshold. Practically, there are not enough observations to calculate 2.5% ES and 2.5% TR from the past 21, 42, 63, and 84 daily returns. Hence, we use the past 100 daily returns to compute 2.5% and 5% ES and TR.

[16] Also note that using $k=5$ or the past 100 days produces very similar results, since 2.5% (5%) TR and ES are computed by using the lowest three (five) observations regardless of whether the past 100 days or the past 105 days (i.e., $k=5$) is used.

Table 8 presents results from the regressions of the value-weighted index return on the lagged TR and ES measures. The slope coefficients on $TR_t^{2.5\%}$ and $TR_t^{5\%}$ are positive and significant, with Newey-West (1987) t-statistics of 1.97 and 1.96, respectively. Similarly, the ES measures, $ES_t^{2.5\%}$ and $ES_t^{5\%}$, have significantly positive slope coefficients, with t-statistics of 2.23 and 2.01, respectively. Indeed, when we exclude macroeconomic variables, the statistical significance goes higher such that $TR_t^{2.5\%}$ and $TR_t^{5\%}$ have t-statistics of 2.56 and 2.55, respectively, and $ES_t^{2.5\%}$ and $ES_t^{5\%}$ have t-statistics of 2.65 and 2.29, respectively (not shown). Overall, the parameter estimates in Table 8 indicate that alternative measures of left-tail risk measures predict the 1-month-ahead market returns almost as well as VaR.

TABLE 8

Alternative Measures of Downside Risk

Table 8 presents results from the time-series regressions of the excess market return on the 2.5% and 5% tail risk ($TR_t^{2.5\%}$ and $TR_t^{5\%}$) and the 2.5% and 5% expected shortfall ($ES_t^{2.5\%}$ and $ES_t^{5\%}$) measures. Here, 2.5% (5%) tail risk is computed as the sum of squared deviations of the lowest 2.5 percentile (5 percentile) of the NYSE/AMEX/Nasdaq index returns from the mean of index returns during the latest 100 days, and 2.5% (5%) expected shortfall is computed as the average of the lowest 2.5 percentile (5 percentile) of NYSE/AMEX/Nasdaq index returns observed during the last 100 days as of the end of month t. In each regression, the first row gives the estimated coefficients. The second row gives the Newey-West (1987) adjusted t-statistics in parentheses. The R^2 values are reported in the last column.

Constant	DR_t	R_t	Dummy	$RREL_t$	$TERM_t$	DEF_t	DP_t	R^2
DR_t = 2.5% Tail Risk								
−0.005	0.956	0.022	−0.114	−0.478	−0.714	3.623	0.291	4.7%
(−0.80)	(1.97)	(0.54)	(−5.56)	(−2.65)	(−2.37)	(2.17)	(1.38)	
DR_t = 5% Tail Risk								
−0.006	0.862	0.022	−0.112	−0.479	−0.715	3.599	0.296	4.7%
(−0.86)	(1.96)	(0.55)	(−5.63)	(−2.65)	(−2.38)	(2.15)	(1.40)	
DR_t = 2.5% Expected Shortfall								
−0.012	0.366	0.028	−0.108	−0.465	−0.723	3.326	0.326	4.9%
(−1.56)	(2.23)	(0.68)	(−6.14)	(−2.54)	(−2.41)	(1.95)	(1.51)	
DR_t = 5% Expected Shortfall								
−0.012	0.414	0.028	−0.102	−0.474	−0.728	3.333	0.326	4.8%
(−1.54)	(2.01)	(0.70)	(−6.16)	(−2.59)	(−2.42)	(1.93)	(1.50)	

C. Risk-Return Tradeoff Over the Long Sample

In this section, as an additional robustness check, we test whether the strong positive relation between downside risk and expected returns persists over the long sample period. Specifically, we use the NYSE/AMEX/Nasdaq index for the sample period of January 1926 to December 2005 and examine the predictive power of VaR.

As presented in Panel A of Table 9, there is a positive and highly significant relation between VaR and excess market return. Specifically, the slope coefficients on VaR have Newey-West (1987) t-statistics in the range of 2.45 to 3.43 for the long sample period of January 1926 to December 2005. Panel B of Table 9 shows that after controlling for the commonly used macroeconomic variables, the strong positive relation remains intact over the sample period of July 1927 to December 2005.[17] Hence, we conclude that the positive downside risk-return tradeoff extends to earlier periods.

[17]The monthly data on 10-year Treasury bonds start from April 1953, and the monthly data on 1-month Treasury bills start from July 1926. Since the data on term spread (defined as the difference

TABLE 9

Downside Risk and Return Tradeoff Over the Long Sample

Table 9 presents results from the time-series regressions of the excess market return on nonparametric Value at Risk (VaR_t) over the long sample period of January 1926 to December 2005 (Panel A) and July 1927 to December 2005 (Panel B). The original VaRs are multiplied by −1 before running our regressions. In each regression, the dependent variable is the 1-month-ahead excess market return, R_{t+1}. The first row gives the estimated coefficients. The second row gives the Newey-West (1987) adjusted t-statistics in parentheses. The R^2 values are reported in the last column.

Panel A. Mean-VaR Tradeoff

k	Constant	VaR_t	R_t	Dummy	R^2
1	0.000	0.845	0.302	−0.051	8.2%
	(−0.07)	(2.86)	(3.11)	(−2.34)	
2	0.003	0.530	0.265	−0.031	7.3%
	(0.57)	(2.45)	(2.79)	(−1.57)	
3	0.001	0.528	0.258	−0.031	7.6%
	(0.27)	(2.95)	(2.79)	(−1.49)	
4	0.001	0.477	0.253	−0.027	7.5%
	(0.34)	(3.43)	(2.75)	(−1.22)	
5	0.004	0.362	0.248	−0.018	7.0%
	(1.00)	(3.35)	(2.68)	(−0.81)	
6	0.004	0.315	0.248	−0.014	6.8%
	(1.23)	(3.10)	(2.66)	(−0.62)	

Panel B. Mean-VaR Tradeoff with Macro Variables

k	Constant	VaR_t	R_t	Dummy	$RREL_t$	DEF_t	DP_t	R^2
1	−0.020	0.707	0.331	−0.021	−0.359	5.304	0.568	10.8%
	(−1.74)	(2.35)	(2.67)	(−0.77)	(−1.65)	(1.40)	(2.11)	
2	−0.019	0.355	0.291	0.001	−0.383	4.567	0.641	9.9%
	(−1.57)	(1.93)	(2.47)	(0.04)	(−1.77)	(1.31)	(2.29)	
3	−0.019	0.373	0.287	−0.001	−0.374	4.526	0.625	10.1%
	(−1.63)	(2.63)	(2.48)	(−0.02)	(−1.71)	(1.30)	(2.24)	
4	−0.019	0.323	0.282	0.003	−0.356	4.492	0.635	10.0%
	(−1.66)	(2.81)	(2.44)	(0.09)	(−1.62)	(1.28)	(2.21)	
5	−0.018	0.218	0.281	0.011	−0.361	4.655	0.662	9.7%
	(−1.57)	(2.26)	(2.41)	(0.35)	(−1.64)	(1.33)	(2.26)	
6	−0.018	0.179	0.281	0.015	−0.363	4.723	0.671	9.6%
	(−1.55)	(1.82)	(2.40)	(0.45)	(−1.65)	(1.36)	(2.27)	

D. Value at Risk Time Aggregates

We also examine the intertemporal relation between VaR and expected return over the 1-month, 2-month, 3-month, 4-month, 5-month, 6-month, and 1-year nonoverlapping intervals. In other words, we examine the parameter estimates from the longer-term time-series regressions of 1-, 2-, 3-, 4-, 5-, 6-, and 12-month-ahead excess return of the value-weighted NYSE/AMEX/Nasdaq index on the VaR computed as the minimum daily index return during the past 1, 2, 3, 4, 5, 6, and 12 months, respectively. For example, when the dependent variable is 4-month-ahead returns, we compute the VaR using the past 4 months of daily data, and hence the regression covers an 8-month period, as demonstrated below:

between the yields on 10-year and 1-month Treasury) start from April 1953, we do not include TERM in our risk-return regressions. Also note that RREL is defined as the difference between the 1-month T-bill rate minus its 12-month backward-moving average, and hence the monthly data on RREL start from July 1927. Therefore, the results in Panel B are based on RREL, DEF, and DP from July 1927 to December 2005.

Estimation Window for VaR: 4 months	Prediction Window for Returns: 4 months

$t = 0$ $t = 4$ $t = 8$

The results in Table 10 indicate a positive and significant relation between VaR and expected returns for 1- to 4-month investment horizons (covering a period of 2 to 8 months), whereas the relation in 5-month, 6-month, and 1-year horizons is weak. Observe that when 1-, 2-, 3-, and 4-month-ahead returns are used as dependent variables, the t-statistic of VaR varies between 2.00 and 3.14. These results suggest that investors care about downside risk for intermediate holding periods.

TABLE 10

Value at Risk Time Aggregates

Table 10 presents results from the long-term time-series regressions of k-month-ahead excess return of the NYSE/AMEX/Nasdaq index on the Value at Risk (VaR) computed as the minimum daily index return during the past k months. Hence, each regression covers $2 \times k$ months of data. Control variables are defined in Table 2. In each regression, the dependent variable is the 1-month-ahead excess market return, R_{t+1}. The first row gives the estimated coefficients. The second row gives the Newey-West (1987) adjusted t-statistics in parentheses. The R^2 values are reported in the last column.

k	Constant	VaR_t	R_t	Dummy	RREL_t	TERM_t	DEF_t	DP_t	R^2
1	−0.012	0.451	0.063	−0.141	−0.467	−0.739	3.381	0.346	4.9%
	(−1.58)	(2.00)	(1.34)	(−4.44)	(−2.57)	(−2.44)	(2.00)	(1.63)	
2	−0.027	0.844	0.027	−0.148	−0.825	−0.784	3.640	0.713	6.1%
	(−1.91)	(3.14)	(0.49)	(−3.80)	(−2.34)	(−1.89)	(1.55)	(1.77)	
3	−0.037	0.951	0.048	−0.119	−1.037	−0.829	3.733	1.063	6.9%
	(−1.82)	(2.82)	(0.70)	(−2.44)	(−2.11)	(−2.01)	(1.03)	(1.87)	
4	−0.039	0.739	0.055	−0.039	−1.239	−0.824	4.284	1.343	6.3%
	(−1.49)	(2.04)	(0.76)	(−0.71)	(−2.00)	(−1.84)	(0.95)	(1.82)	
5	−0.037	0.534	0.044	−0.035	−1.537	−0.993	10.467	1.571	7.1%
	(−1.23)	(1.60)	(0.55)	(−0.68)	(−2.09)	(−1.83)	(2.21)	(1.77)	
6	−0.029	0.154	0.004	0.012	−2.012	−1.657	12.461	1.792	7.7%
	(−0.83)	(0.50)	(0.04)	(0.25)	(2.33)	(−2.47)	(2.32)	(1.74)	
12	−0.042	0.258	−0.044	0.023	−3.573	−2.028	15.293	3.052	11.2%
	(−0.57)	(0.58)	(−0.36)	(0.36)	(2.56)	(−2.48)	(2.19)	(1.63)	

VII. Comparing Value at Risk with Variance

A. Various Measures of Realized Variance

As mentioned earlier, theoretical models suggest a positive relation between conditional mean and variance of returns for the aggregate stock market. One of the most commonly used estimators of conditional variance is the sum of squared daily returns over the previous month (see French et al. (1987)). Although this measure of market variance has been used extensively in tests of risk-return tradeoff, there is no evidence of a positive and significant relation between this measure of conditional variance and expected returns.

Harrison and Zhang (1999) examine the time-series relation between expected returns and conditional volatility over different holding periods and find a significantly positive risk-return tradeoff at long holding intervals (such as 1 to 2 years), which does not exist at short holding periods (such as 1 month).

Similar to the original findings of Harrison and Zhang (1999), Ghysels et al. (2005) present evidence on the risk-return tradeoff. Like French et al. (1987), they use the rolling window approach and the sum of squared daily returns as a proxy for the monthly conditional variance. Additionally, they argue that since the realized variance is very persistent, it ought to be a good proxy for the conditional variance. On the other hand, they point out that it is not clear why the researchers should confine themselves to using data from the last month only to estimate the conditional variance. Therefore, they use a larger window size (from 1 to 6 months) when they sum the past squared returns to acquire the conditional variance measure. Interestingly, this choice has a tremendous impact on the significance of the risk-return tradeoff. Ghysels et al. (2005) report that the variance measures that are computed using the daily returns over the previous 3 to 6 months significantly forecast the market return.

These recent findings are important for us, because we would like to compare our measure of downside risk with the traditional risk measures that are shown to have a statistically significant predictive power for the expected market returns. Thus, in light of these recent findings, we have to focus on realized variance that is computed using a window size larger than 1 month because the realized variance in the previous month is not a significant predictor of expected returns.

In Table 11, we reexamine the findings of Ghysels et al. (2005) on the rolling window estimates. We assume that each month has 21 trading days and compute the realized variance as the sum of squared daily returns on the value-weighted NYSE/AMEX/Nasdaq index plus an adjustment term for the first-order serial correlation in daily returns. We generate different variance measures for horizons of 1 to 6 months. Panel A of Table 11 presents the parameter estimates from the regressions of excess value-weighted market return on the lagged realized variance, lagged market return, and the dummy for the October 1987 crash. The first column shows the number of months used in the estimation of the conditional variance proxy. Similarly to the very early literature, we find that the realized variance in the previous month has no forecasting power, but from month 2 through month 6, realized variance is a significant forecaster of market returns. In Panel B of Table 11, when we control for the macroeconomic variables, the statistical significance of the realized variance reduces, except for the 3- and 4-month estimation window.

B. Value at Risk versus Realized Variance

We have so far shown that there is a significant relation between VaR and expected returns. However, another objective of this paper is to compare the predictive power of VaR with the predictive power of traditional risk proxies that are shown to forecast market returns. Table 12 compares the relative performance of various VaRs and realized variance measures computed over different horizons in predictive regressions. For 1- to 6-month horizons, we compute VaR and variance measures and use them in the same regressions. For example, at 2 months, we compute VaR as the minimum return observed during the past 42 days and compute the variance as the sum of squared daily returns during the past 42 days plus

TABLE 11

Various Measures of Realized Variance

Table 11 shows estimates of the risk-return tradeoff with the rolling window estimators of realized variance. For each horizon, the realized variance is computed as the sum of the squared daily returns on the value-weighted NYSE/AMEX/Nasdaq index plus an adjustment term for the first-order serial correlation in daily returns. The first column in each panel shows the number of months (k) used to compute the variance. The dependent variable is the 1-month-ahead excess market return, R_{t+1}. Panel A presents the parameter estimates without the control variables, and Panel B presents the parameter estimates with the control variables. In each regression, the first row gives the estimated coefficients. The second row gives the Newey-West (1987) adjusted t-statistics in parentheses. The R^2 values are reported in the last column.

Panel A. Mean-Variance Tradeoff

k	Constant	σ_t^2	R_t	Dummy	R^2
1	0.004	0.462	0.040	−0.101	0.8%
	(1.60)	(0.49)	(0.98)	(−1.67)	
2	0.002	0.684	0.044	−0.115	1.3%
	(0.88)	(1.70)	(1.05)	(−4.39)	
3	0.001	0.627	0.040	−0.112	1.7%
	(0.51)	(2.64)	(0.99)	(−6.09)	
4	0.001	0.516	0.035	−0.106	1.8%
	(0.34)	(2.70)	(0.88)	(−6.24)	
5	0.001	0.342	0.034	−0.094	1.4%
	(0.55)	(1.99)	(0.84)	(−5.85)	
6	0.001	0.270	0.035	−0.090	1.3%
	(0.61)	(1.71)	(0.87)	(−5.83)	

Panel B. Mean-Variance Tradeoff with Macro Variables

k	Constant	σ_t^2	R_t	Dummy	$RREL_t$	$TERM_t$	DEF_t	DP_t	R^2
1	−0.003	−0.060	0.021	−0.076	−0.517	−0.751	3.849	0.283	4.2%
	(−0.48)	(−0.06)	(0.50)	(−1.21)	(−2.84)	(−2.51)	(2.20)	(1.33)	
2	−0.006	0.549	0.031	−0.114	−0.492	−0.781	3.284	0.289	4.5%
	(−0.88)	(1.23)	(0.73)	(−3.89)	(−2.77)	(−2.57)	(1.82)	(1.36)	
3	−0.007	0.497	0.028	−0.111	−0.468	−0.754	3.240	0.290	4.7%
	(−0.98)	(2.04)	(0.69)	(−5.84)	(−2.61)	(−2.51)	(1.83)	(1.36)	
4	−0.007	0.402	0.025	−0.106	−0.459	−0.751	3.203	0.292	4.8%
	(−1.02)	(2.04)	(0.61)	(−5.93)	(−2.52)	(−2.51)	(1.79)	(1.36)	
5	−0.006	0.225	0.023	−0.093	−0.473	−0.718	3.515	0.280	4.4%
	(−0.85)	(1.30)	(0.56)	(−5.58)	(−2.60)	(−2.39)	(1.99)	(1.31)	
6	−0.006	0.169	0.023	−0.090	−0.478	−0.717	3.612	0.280	4.4%
	(−0.82)	(1.11)	(0.56)	(−5.71)	(−2.61)	(−2.39)	(2.07)	(1.32)	

the autocorrelation adjustment term. We further include the control variables to run a full specification.

Panel A of Table 12 shows that, at all horizons, the nonparametric VaR measure has a positive and significant coefficient estimate. The Newey-West (1987) t-statistics are in the range of 2.10 to 3.49 when the past 1 to 5 months of daily returns are used in nonparametric VaR calculations. For the 6-month horizon, the slope coefficient on VaR has a t-statistic of 1.91 with a p-value of 5.61%. On the other hand, the coefficients on rolling window estimators of variance are negative and statistically insignificant. Therefore, measuring the variance by using a window size larger than 1 month has a substantial effect on the risk-return tradeoff, but that impact is captured by VaR. We conclude that VaR is not only a good measure of downside risk that is related to expected returns, it also captures information about expected returns that cannot be explained by the traditional measure of market risk, realized variance, even if it is computed by using a larger number of observations.

Panel B of Table 12 confirms these findings based on the parametric VaR. It is clear that parametric VaR also outperforms the rolling window estimators of

TABLE 12

Comparing Value at Risk with Realized Variance

Table 12 compares the relative performance of rolling window estimates of the Value at Risk (VaR) and realized variance in predicting future market returns. Each month is assumed to have 21 trading days. The first column in each panel shows the number of months (k) used to compute the VaR and variance. In Panel A, for each horizon the nonparametric VaR (VaR_t) is computed as the lowest daily returns on the NYSE/AMEX/Nasdaq index. Similarly, for each horizon, the realized variance is computed as the sum of squared daily returns on the NYSE/AMEX/Nasdaq index plus an autocorrelation adjustment term. In Panel B, for each horizon the parametric VaR (VaR_t^P) is computed as the lower tail of the skewed t density. In each regression, the dependent variable is the 1-month-ahead excess market return, R_{t+1}. The first row gives the estimated coefficients. The second row gives the Newey-West (1987) adjusted t-statistics in parentheses. The R^2 values are reported in the last column.

k	Constant	VaR_t	σ_t^2	R_t	Dummy	RREL_t	TERM_t	DEF_t	DP_t	R^2
Panel A. Relative Performance of Nonparametric VaR and Variance										
1	−0.015	0.902	−2.453	0.077	−0.049	−0.473	−0.721	3.956	0.386	5.6%
	(−2.03)	(3.19)	(−2.10)	(1.70)	(−0.85)	(−2.58)	(−2.48)	(2.29)	(1.86)	
2	−0.018	0.824	−1.115	0.054	−0.126	−0.468	−0.709	3.429	0.401	6.2%
	(−2.24)	(3.49)	(−1.79)	(1.24)	(−5.51)	(−2.58)	(−2.41)	(2.06)	(1.83)	
3	−0.015	0.580	−0.532	0.035	−0.130	−0.477	−0.742	3.414	0.380	5.8%
	(−1.89)	(2.42)	(−1.09)	(0.86)	(−7.25)	(−2.63)	(−2.50)	(2.07)	(1.72)	
4	−0.015	0.560	−0.481	0.033	−0.131	−0.464	−0.758	3.303	0.386	6.0%
	(−1.94)	(2.51)	(−1.08)	(0.81)	(−7.70)	(−2.54)	(−2.58)	(1.96)	(1.74)	
5	−0.012	0.412	−0.378	0.027	−0.116	−0.482	−0.711	3.650	0.356	5.3%
	(−1.57)	(2.10)	(−1.01)	(0.66)	(−6.67)	(−2.64)	(−2.40)	(2.16)	(1.64)	
6	−0.011	0.346	−0.301	0.026	−0.110	−0.489	−0.718	3.728	0.350	5.1%
	(−1.43)	(1.91)	(−0.92)	(0.64)	(−6.57)	(−2.67)	(−2.42)	(2.21)	(1.61)	
Panel B. Relative Performance of Parametric VaR and Variance										
1	−0.016	0.842	−2.523	0.076	−0.046	−0.480	−0.729	3.969	0.394	5.5%
	(−2.14)	(3.17)	(−2.07)	(1.70)	(−0.78)	(−2.61)	(−2.54)	(2.29)	(1.91)	
2	−0.018	0.803	−1.027	0.053	−0.126	−0.489	−0.735	3.410	0.402	6.2%
	(−2.18)	(3.22)	(−1.53)	(1.23)	(−5.33)	(−2.65)	(−2.48)	(2.04)	(1.81)	
3	−0.014	0.548	−0.392	0.037	−0.129	−0.493	−0.760	3.339	0.370	5.8%
	(−1.78)	(2.10)	(−0.77)	(0.89)	(−7.26)	(−2.67)	(−2.54)	(1.99)	(1.67)	
4	−0.015	0.618	−0.458	0.036	−0.134	−0.491	−0.783	3.285	0.382	6.0%
	(−1.96)	(2.46)	(−1.01)	(0.88)	(−7.71)	(−2.63)	(−2.65)	(1.96)	(1.74)	
5	−0.011	0.430	−0.320	0.028	−0.117	−0.503	−0.729	3.647	0.341	5.3%
	(−1.50)	(1.86)	(−0.83)	(0.68)	(−6.55)	(−2.71)	(−2.45)	(2.15)	(1.59)	
6	−0.010	0.362	−0.252	0.026	−0.112	−0.508	−0.732	3.723	0.330	5.1%
	(−1.35)	(1.70)	(−0.76)	(0.64)	(−6.40)	(−2.72)	(−2.46)	(2.19)	(1.54)	

variance for all horizons considered in the paper. The slope coefficients on parametric VaR have t-statistics ranging from 2.10 to 3.22 when the past 1 to 4 months of daily returns are used in parametric VaR calculations. For 5- and 6-month estimation windows, the parametric VaR has a positive but marginally significant coefficient estimate. A notable point in Table 12 is that the coefficient estimates on alternative measures of variance are negative and statistically insignificant.

VIII. Conclusion

We examine the intertemporal relation between downside risk and expected stock returns. We use Value at Risk (VaR) as a measure of downside risk and find a positive and significant relation between VaR and expected return on the market. Moreover, we generate alternative measures of VaR based on the past 1 to 6 months of daily data, and we show that there is a significantly positive relation between VaR and expected market return for all horizons considered in the paper. Finally, we test the relative performance of various VaRs and realized variance measures computed over different horizons in predictive regressions. The results indicate that VaR wins convincingly even when it is compared to the conditional variance proxies that have significant predictive power for market returns. These findings are robust across different measures of market return, loss probability levels, and after controlling for macroeconomic variables associated with business cycle fluctuations.

If downside risk is an important determinant of expected returns, we expect other proxies of downside risk to perform well in predictive regressions too. Therefore, we use expected shortfall and tail risk, both of which inform us about the left tail of the return distribution, as alternative measures of downside risk. We show that, regardless of the left-tail measure we use, our qualitative results from predictive regressions remain intact.

The results provide strong evidence that there exists a positive and significant relation between downside risk and expected returns. Our findings also suggest that rare large moves in the market or relatively infrequent return observations can be interpreted as signals, whereas the frequent small fluctuations can be viewed as noise, which does not have the power to explain time-series variation in excess market returns.

References

Ang, A.; J. Chen; and Y. Xing. "Downside Risk." *Review of Financial Studies*, 19 (2006), 1191–1239.

Arditti, F. D. "Risk and the Required Return on Equity." *Journal of Finance*, 22 (1967), 19–36.

Arditti, F. D., and H. Levy. "Portfolio Efficiency Analysis in Three Moments: The Multi-Period Case." *Journal of Finance*, 30 (1975), 797–809.

Artzner, P.; F. Delbaen; J.-M. Eber; and D. Heath. "Coherent Measures of Risk." *Mathematical Finance*, 9 (1999), 203–228.

Arzac, E. R., and V. S. Bawa. "Portfolio Choice and Equilibrium in Capital Markets with Safety-First Investors." *Journal of Financial Economics*, 4 (1977), 277–288.

Bali, T. G. "An Extreme Value Approach to Estimating Volatility and Value at Risk." *Journal of Business*, 76 (2003), 83–108.

Bali, T. G. "The Intertemporal Relation between Expected Returns and Risk." *Journal of Financial Economics*, 87 (2008), 101–131.

Bali, T. G.; N. Cakici; X. Yan; and Z. Zhang. "Does Idiosyncratic Risk Really Matter?" *Journal of Finance*, 60 (2005), 905–929.

Baumol, W. J. "An Expected Gain-Confidence Limit Criterion for Portfolio Selection." *Management Science*, 10 (1963), 174–182.

Berk, J. B. "Necessary Conditions for the CAPM." *Journal of Economic Theory*, 73 (1997), 245–257.

Black, F., and M. Scholes. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy*, 81 (1973), 637–654.

Campbell, J. Y. "Stock Returns and the Term Structure." *Journal of Financial Economics*, 18 (1987), 373–399.

Carr, P.; H. Geman; D. B Madan; and M. Yor. "The Fine Structure of Asset Returns: An Empirical Investigation." *Journal of Business*, 75 (2002), 305–332.

Chamberlain, G. "A Characterization of the Distributions that Imply Mean-Variance Utility Functions." *Journal of Economic Theory*, 29 (1983), 185–201.

Chen, J.; H. Hong; and J. Stein. "Forecasting Crashes: Trading Volume, Past Returns, and Conditional Skewness in Stock Prices." *Journal of Financial Economics*, 61 (2001), 345–381.

Cornish, E. A., and R. A. Fisher. "Moments and Cumulants in the Specification of Distributions," *Extrait de la Revue de l'Institute International de Statistique*, 4 (1937), 1–14. Reprinted in Fisher, R. A., *Contributions to Mathematical Statistics*. New York: Wiley (1950).

Cox, J. C., and S. A. Ross. "The Valuation of Options for Alternative Stochastic Processes." *Journal of Financial Economics*, 3 (1976), 145–166.

Dittmar, R. F. "Nonlinear Pricing Kernels, Kurtosis Preference, and Evidence from the Cross Section of Equity Returns." *Journal of Finance*, 57 (2002), 369–403.

Fama, E. F., and K. R. French. "Dividend Yields and Expected Stock Returns." *Journal of Financial Economics,* 22 (1988), 3–25.
Fama, E. F., and K. R. French. "Business Conditions and Expected Returns on Stocks and Bonds." *Journal of Financial Economics,* 25 (1989), 23–49.
Fama, E. F., and J. D. MacBeth. "Risk, Return, and Equilibrium: Empirical Tests." *Journal of Political Economy,* 81 (1973), 607–636.
Ferson, W. E., and C. Harvey. "The Variation of Economic Risk Premiums." *Journal of Political Economy,* 99 (1991), 385–415.
French, K. R.; G. W. Schwert; and R. F. Stambaugh. "Expected Stock Returns and Volatility." *Journal of Financial Economics,* 19 (1987), 3–29.
Ghysels, E.; P. Santa-Clara; and R. Valkanov. "There is a Risk-Return Trade-Off After All." *Journal of Financial Economics,* 76 (2005), 509–548.
Guo, H., and R. Whitelaw. "Uncovering the Risk-Return Relation in the Stock Market." *Journal of Finance,* 61 (2006), 1433–1463.
Hansen, B. E. "Autoregressive Conditional Density Estimation." *International Economic Review,* 35 (1994), 705–730.
Harrison, P., and H. H. Zhang. "An Investigation of the Risk and Return Relation at Long Horizons." *Review of Economics and Statistics,* 81 (1999), 399–408.
Harvey, C. R., and A. Siddique. "Conditional Skewness in Asset Pricing Tests." *Journal of Finance,* 55 (2000), 1263–1295.
Jarque, C. M., and A. K. Bera. "Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals." *Economics Letters,* 6 (1980), 255–259.
Kimball, M. S. "Standard Risk Aversion." *Econometrica,* 61 (1993), 589–611.
Kraus, A., and R. H. Litzenberger. "Skewness Preference and the Valuation of Risk Assets." *Journal of Finance,* 31 (1976), 1085–1100.
Levy, H., and H. Markowitz. "Approximating Expected Utility by a Function of Mean and Variance." *American Economic Review,* 69 (1979), 308–317.
Levy, H., and M. Sarnat. "Safety First—An Expected Utility Principle." *Journal of Financial and Quantitative Analysis,* 7 (1972), 1829–1834.
Longin, F. M. "From Value at Risk to Stress Testing: The Extreme Value Approach." *Journal of Banking and Finance,* 24 (2000), 1097–1130.
Markowitz, H. "Portfolio Selection." *Journal of Finance,* 7 (1952), 77–91.
Markowitz, H. *Portfolio Selection: Efficient Diversification of Investments*, 2nd ed. Malden, MA: Blackwell (1959).
Merton, R. C. "An Intertemporal Capital Asset Pricing Model." *Econometrica,* 41 (1973), 867–887.
Merton, R. C. "Option Pricing When Underlying Stock Returns are Discontinuous." *Journal of Financial Economics,* 3 (1976), 125–144.
Merton, R. C. "On Estimating the Expected Return on the Market: An Exploratory Investigation." *Journal of Financial Economics,* 8 (1980), 323–361.
Neftci, S. N. "Value at Risk Calculations, Extreme Events, and Tail Estimation." *Journal of Derivatives,* 7 (2000), 23–37.
Newey, W. K., and K. D. West. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica,* 55 (1987), 703–708.
Pratt, J. W., and R. J. Zeckhauser. "Proper Risk Aversion." *Econometrica,* 55 (1987), 143–154.
Roy, A. D. "Safety First and the Holding of Assets." *Econometrica,* 20 (1952), 431–449.
Scott, R. C., and P. A. Horvath. "On the Direction of Preference for Moments of Higher Order than the Variance." *Journal of Finance,* 35 (1980), 915–919.
Tallarini, T. D., and H. H. Zhang. "External Habit and the Cyclicality of Expected Stock Returns." *Journal of Business,* 78 (2005), 1023–1048.