Helsinki Center of Economic Research

Discussion Papers

Testing for Common Costs in the City of Helsinki Bus Transit Auctions

Janne Tukiainen
University of Helsinki, RUESG and HECER

Discussion Paper No. 157
April 2007

ISSN 1795-0562

HECER – Helsinki Center of Economic Research, P.O. Box 17 (Arkadiankatu 7), FI-00014 University of Helsinki, FINLAND, Tel +358-9-191-28780, Fax +358-9-191-28781,
E-mail info-hecer@helsinki.fi, Internet www.hecer.fi

HECER
Discussion Paper No. 157

Testing for Common Costs in the City of Helsinki Bus Transit Auctions*

Abstract

I study the effects of competition in the City of Helsinki bus transit markets by conducting two tests for common costs in these first-price sealed-bid procurement auctions. I introduce pooling and bidder asymmetry to the tests. I also show the need for robustness checks for some arbitrary choices in these tests. The information environment seems to be that of common costs. The bus companies that have garages close to the contracted routes are more influenced by the common elements than those whose garages are further away. More competition does not necessary lower procurement costs and the City should not necessarily implement costly policies to induce more competition. Also the recent merger of the two public companies cannot be criticized from a competition perspective.

JEL Classification: C12,C14,C15,D44,L52,L92

Keywords: Bidder asymmetry, bus transit, common values, first-price auctions, nonparametric testing and private values.

Janne Tukiainen

Department of Economics
University of Helsinki
P.O. Box 17 (Arkadiankatu 7)
FI-00014 University of Helsinki
FINLAND

e-mail: janne.tukiainen@helsinki.fi

* I would like to thank Tero Anttila, Estelle Cantillon, Philip Haile, Heikki Kauppi, Tommi Kyntölä, Heikki A. Loikkanen, Paavo Miettinen, Heikki Pursiainen, Kalevi Saarelainen, Otto Toivanen and Jouko Verho for help and comments. Seminar participants at several HECER and FDPE seminars, ESEM-2006 and ECARES provided helpful insights. All the possible mistakes are mine. I gratefully aknowledge financial support from FDPE, Yrjö Jahnsson Foundation, Finnish Cultural Foundation, the City of Helsinki and Kunnallisalan Kehittämissäätiö.

1 Introduction

In this paper, I study two policy questions concerning the bus transit market in the City of Helsinki. Both questions are viewed from the perspective of the effects of competition. The first is whether the merger of the City owned private company Suomen Turistiauto Oy and the Helsinki City Transport's (HKL) bus transport department, HKL-bussiliikenne, in 1.1.2005 was good policy. The second is whether, as planned, the City should spend resources to induce more competition in this market. One possibility that has been considered by the procurement officials is to build new City owned garages and rent these to new entrants.

A central issue in an auction setting when assessing the effects of competition, i.e. number of bidders, on procurement costs is whether the bidders operate in a private or a common cost environment. Common costs refer to the situation where the relevant information about the costs of providing the contracted services is dispersed among the bidders. In that case bidders would update their beliefs about these costs if they learned their competitors' signals on these costs. Private costs refer to the situation where the bidders are only interested in their own signals when evaluating the true costs. This distinction is called the information paradigm. The model by Milgrom and Weber (1982) (denoted MW) suggests that the effects of competition may change with the information paradigm. This is due to a phenomenon known as a winner's curse. The winner's curse refers to a situation where bidders bid in a common costs environment only according to their own cost estimate. With unbiased estimates and symmetric bidders, the bidder who underestimates his costs most wins the auctions and may receive a negative payoff. The expected amount of underestimation increases with the number of bidders. Rational bidders take this into account and thus may even bid more as competition increases. Hong and Shum (2002) find empirical evidence of this counterintuitive effect. Strategic behavior implies that bidders bid less when the number of bidders increases. With private costs only this strategic component is in play, whereas in common cost setting both of these factors matter and thus then the effect of competition is not clear. This distinction forms the basis of this study.

I perform two tests developed by Haile at al. (2003) (denoted HHS) for common costs. The setting a is first-price sealed-bid procurement auction. The main contribution of this paper is the empirical application and its policy implications. Novelty is achieved by extending the testing framework by allowing the bidders to be asymmetric. Another extension is that I use pooling to meet the data requirements. Another aspect of interest is showing the need for robustness checks for some arbitrary choices in HHS testing procedures. Also the discussion on potential misspecification pitfalls is of interest beyond this application. To my knowledge this is the second study that applies the HHS methodology. In the other (Shneyerov 2003), only the means[1] test is used.

According to the Confederation of Finnish Industries (www.ek.fi 2005) public procurement in Finland amounts to about 20 billion euros annually. This is about 15 % of GDP. In the European Union the share of public procurement is about 16 % of GDP. Despite it's importance, procurement from the auction perspective has not been studied with Finnish data before. The nationality of data is important from the policy point of view. One important goal of the procurement officials is to increase competition (Saarelainen 2002). The increased competition is meant to lower procurement costs and increase the quality of services. Theoretical and empirical results quoted above imply that this is not necessarily the case if bidders have common costs. Thus different procurement rules should be implemented for different markets. For example the EU law of open invitation to all procurements is perhaps not optimal. Besides the effects of competition, HHS tests can be used to determine the optimal auction format. MW show that if the winning bidder's payoff depends on the preferences of others, implying common costs, ascending auctions generate lower procurement costs than other auction types.

In many econometric studies the choice of the information paradigm is based on intuition. To simplify, in situations where bidders input-output efficiency is dominant and the only uncertainty is about other competitors' efficiencies, one should assume private costs, whereas in situations where there exists uncertainty on common elements like technological development, one should assume common costs. In their study of London bus transit auctions Cantillon and Pesendorfer (2006b) assume and argue convincingly for private costs. Moreover, in the auctions studied in this paper, the winners are compensated for the changes in most input prices following the Statistics Finland bus transit cost index. This reduces uncertainty about the future and thus the relative importance of common elements. Therefore the theoretical (also technical) null hypothesis in this study is that of private costs. This null is however rejected in most tests. This surprising result can also be due to model misspecification.

In the symmetric case the null hypothesis of the private costs paradigm is rejected in most specifica-

[1] Actually medians in that application.

tions. Therefore common costs play a role in this market. When allowing for asymmetry I find that these common costs are driven by the bus companies that have garages close to the contracted routes. This result has an important policy implication. If planners choose between two locations, from which the overall operation costs are equal, they should choose to build the public garage further away. Moreover it is not clear whether this policy should be implemented at all, as the benefits of more competition are not clear. Another implication is that the merger of the two public bus companies cannot be criticized on the basis of it reducing competition and thus increasing procurement costs.

In practise the HHS tests are conducted in the following way. First the bidder's expected costs conditional on winning the auction are estimated using observed bids with structural nonparametric methods. This follows Li et al. (2002) (denoted LPV) for the symmetric case and Campo et al. (2003) (denoted CPV) for the asymmetric case. In the common costs setting these costs are increasing in the number of bidders when bidders are symmetric. With asymmetric bidders and common costs this relation is not clear. With private costs these costs are invariant in number of bidders with both symmetric and asymmetric bidders. Therefore two tests for stochastic dominance between the cost distributions for different numbers of bidders are conducted. In the asymmetric case also the types of the bidders in addition to their number have to be taken into account. The interpretation of the null hypotheses also differs in these two cases. The first test compares quantile trimmed means and utilizes bootstrapping procedures. The second is a modified Kolmogorov-Smirnov test based on subsampling.

Besides the policy questions, the assumption on the information paradigm has severe implications on the analysis of a given market. The HHS test is therefore also useful in determining which methods should be applied with a given data. HHS mention three arguments why the nonparametric nature of their tests avoids the problems with previous suggestions (e.g. Paarsch (1992), Sareen (1999), Gilley and Karels (1981) and Hendricks at al. (2003)). The first problem is confounding testing for the information paradigm with testing for parametric assumptions. The second is avoiding to have to base the test on solely bids. The third is not having to test for a particular form of private and common values. Pinkse and Tan (2005) suggest another possible alternative to testing using winning bids, but they do not develop the test explicitly.

Although auction theory has been fairly well established for many decades (e.g. Krishna (2002) or Milgrom (2004)), empirical work is new in comparison. Earlier empirical literature is surveyed by Laffont

(1997) and Hendricks and Paarsch (1995). Early studies utilized mostly reduced form econometric models. Perrigne and Vuong (1999) survey earlier methods for structural analysis of first-price auctions. Kagel (1995) surveys experimental work on auctions. The structural empirical analysis of auctions has developed rapidly in recent years. This literature is surveyed by Hendricks and Porter (2005), Paarsch and Hong (2006) and Athey and Haile (2005). Paarsch (1992), Donald and Paarsch (1993,1996) and Laffont et al. (1995) developed the first parametric structural methods. Elaykime et al. (1994) and Guerre et al. (2000) have developed nonparametric structural methods, which LPV extended to affiliated values. Until LPV the analysis had been confined to either common value or private value models. CPV made the extension for identifying the asymmetric affiliated case. Bus transit data have been used previously in the empirical auction literature by Cantillon and Pesendorfer (2006a, 2006b) who examine the efficiency of package (or combination) bidding. Previous studies on the Helsinki bus transit market have concentrated on the effects of the change of regime that happened in 1997 when the city started to tender the bus transit services instead of buying them through negotiations. For example Valkama and Finkkilä (2003) have studied economic issues like the changes in the bus companies' economic performance. Haatainen (2003) studied the effects on the bus companies' personnel.

2 The theoretical model and the estimation strategy

Usually a researcher observes the bids submitted in auctions. This is unfortunately not enough to directly infer under which information paradigm the bidders operate. With common costs the winner's curse and strategic behavior have opposite effects on the bids as the amount of competition changes. Pinkse and Tan (2005) showed that in affiliated private values first price auctions the affiliation effect causes the same response to increased competition as the winner's curse. Therefore, to be able to test for the paradigm we need to estimate the bidders' costs or their conditional expectations on costs that correspond to the bids we observe. In that way the strategic behavior is controlled for. Inference on the paradigm can be conducted because these costs react to competition differently depending on the paradigm.

I will present only the basic idea of the model and estimation here. For more one should refer to MW, LPV, CPV and HHS. In appendix 1 there is a longer walkthrough of the estimation for those not familiar with the literature. It makes it possible to replicate what I have done. The tests in HHS are based on comparing the distributions of bidders' expected costs conditional on winning the auction for different numbers of bidders. For bidder i this cost is $v_n(X_i, Y_i) = E[C_i \mid X_i, Y_i]$, where X_i is a private signal of bidder i and Y_i is lowest signal among the $(n-1)$ other bidders. The estimation of these costs is done following LPV and CPV. Estimated costs are called pseudo-costs.

The models are based on the assumptions that bidders use symmetric Bayesian Nash equilibrium strategies and maximize their expected payoff conditional on their signal. The first order condition is sufficient for estimation and actual strategies need not to be solved. Identification is based on the monotonicity of bidding strategies $s_n(x)$ with respect to signals on costs x. Pseudo-costs can be estimated from bid information alone with nonparametric methods when all bids are observed. Equation (2.1) is the essential first order condition for equilibrium for bidder i in symmetric n bidder auction. This well known condition is from MW. $F_{Y_i,n}$ is the distribution of Y_i conditional on bidder's own signal being x and $f_{Y_i,n}$ is the corresponding density. See appendix 1 for the asymmetric equivalent.

$$(2.1) \qquad s_n\acute{}(x) = [s_n(x) - v_n(x,x)]\frac{f_{Y_i,n}(x|x)}{1-F_{Y_i,n}(x|x)}.$$

3 Testing

This section presents the principle of the HHS tests and the asymmetric modification. I show shortly in appendix 2 how all the necessary calculations are made in practise. For more detail refer to HHS. HHS base their tests on a formal definition of private and common costs.

Definition 1. Bidders have private costs if and only if $E[C_i \mid X_1, ..., X_n] = E[C_i \mid X_i]$ and bidders have common costs if and only if $E[C_i \mid X_1, ..., X_n]$ strictly increases in X_j for $j \neq i$.

This definition of common costs incorporates a wide range of models with a common cost component, not just the special case of pure common costs (HHS). HHS prove that definition 1 leads to the following essential theorem.

Theorem 1. With private costs $v_n(x,x)$ is invariant to n for all x but with common costs *and symmetric bidders* strictly decreasing in n for all x.

In this form theorem 1 was presented in Athey and Haile (2005, p.94). HHS then show that this implies

Corollary 1. Under the private cost hypothesis
$F_{v,\underline{n}}(v) = F_{v,\underline{n}+1}(v) = ... = F_{v,\bar{n}}(v)\ \forall v.$
Under the common cost hypothesis and with symmetric bidders
$F_{v,\underline{n}}(v) > F_{v,\underline{n}+1}(v) > ... > F_{v,\bar{n}}(v)\ \forall v.$

This forms the test hypothesis in the symmetric case. Assuming two types of bidders and letting n_1 denote the number of bidders of type 1 and n_0 denote the number of type 0 bidders, we can see from Theorem 1 also that

Corollary 2. Under the private cost hypothesis and with asymmetric bidders
$F_{v,\underline{n}_0,\underline{n}_1}(v) = F_{v,\underline{n}_0+1,\underline{n}_1}(v) = F_{v,\underline{n}_0+1,\underline{n}_1+1}(v) = ... = F_{v,\bar{n}_0,\bar{n}_1}(v)\ \forall v.$
Under the common cost hypothesis and with asymmetric bidders
$F_{v,\underline{n}_0,\underline{n}_1}(v) = F_{v,\underline{n}_0+1,\underline{n}_1}(v) = F_{v,\underline{n}_0+1,\underline{n}_1+1}(v) = ... = F_{v,\bar{n}_0,\bar{n}_1}(v)\ \forall v.$ *or*
$F_{v,\underline{n}_0,\underline{n}_1}(v) > F_{v,\underline{n}_0+1,\underline{n}_1}(v) > F_{v,\underline{n}_0+1,\underline{n}_1+1}(v) > ... > F_{v,\bar{n}_0,\bar{n}_1}(v)\ \forall v.$

Thus if we observe equal distributions the information paradigm is unknown but if we observe unequal distributions the environment must be common. This forms the test hypothesis in the asymmetric case. This is not the entire partition of the relation set but in practise the relation in the other direction should not be observed. Two sided tests should be conducted if counterintuitive direction is observed. The competition in the asymmetric case must be defined correctly. There must be at least the same amount of one type of bidders and more of the other type of bidders to be able to compare distributions with maintaining this interpretation.

HHS then propose to use the empirical distribution function $\hat{F}_{v,n}(y) = \frac{1}{T_n \times n} \sum_{t=1}^{T} \sum_{i=1}^{n} 1\{\hat{v}_{it} < y, n_t = n\}$ to conduct the tests. They explain the difficulties involved in tests that use nonparametrically estimated pseudo-costs $\hat{v}_{it}$. HHS suggest the use of two tests. These tests are based on testing for stochastic dominance between different empirical distributions of the estimated pseudo-costs. The null hypothesis is that of equal distributions. The idea of their first test is to compare the distributions horizontally using quantile trimmed means. They use block bootstrapping to calculate variances. As Athey and Haile (2005) summarize it, the test is an adaptation of a standard multivariate one-sided likelihood-ratio test by Bartholomew (1959). Their second and their preferred test is a generalization of the Kolmogorov-Smirnov test. The idea is to compare the distributions vertically using the sum of maximum distances as test statistic. They normalize their test statistic to be able to use subsampling to estimate critical values. HHS also suggest recentering.

4 Market, data and misspecification

This section has three objectives. First, I describe the Helsinki metropolitan area bus transit market. Second, I present descriptive summaries of the data. Third, and most importantly, I discuss the problems of applying the HHS methodology to this particular data set. These problems are the small size of the data set and the different rules used in the actual auction from the one assumed in the methods.

4.1 The Helsinki bus transit market

The Helsinki metropolitan area consists of four different cities, Helsinki, Espoo, Vantaa and Kauniainen. The population of Helsinki makes 56 percent of the total population of one million and it is the most densely populated one. Kauniainen is practically a small suburb of Espoo. The Helsinki metropolitan area bus transit market represents about 270 routes serving about 172 million passengers per year. It is valued at about 300 million euros per year. The market can be divided into four parts: regional bus transit, and the intra-city bus transit in Espoo, Helsinki and Vantaa. The Helsinki Metropolitan Area Council (YTV) arranges the regional tenders as well as the tenders in Espoo and Vantaa. The City of Helsinki arranges its own tenders. Both use similar rules and have for example standardized the vehicle requirements in most respects. Intra-city traffic of Helsinki consists of 86 routes. Its share of the total market was about 58 percent of passengers, 41 percent of bus kilometers and 59 percent of expenditures in year 2000. The first tender was awarded in 1994 for regional transport services. The City of Helsinki arranged its first tender in 1997, the City of Espoo tendered all its services in one go in 1998 and the City of Vantaa did the same in 1999. Introducing tendering in the place of buying the services through negotiations reduced the procurement costs by 15,2 % in Helsinki and more in the other three markets (Saarelainen 2004). It is not clear what is the welfare effect of this regime change. These savings can be due to genuine efficiency improvements and getting rid of monopoly rents, or they can be just transfers from employees (after the regime change media reported depreciating working conditions) and entrepreneurs to the public sector.

The data set used in this study includes only the City of Helsinki tenders. The data consists of 64 auctions. The small number of observations causes problems with statistical inference, as both the nonparametric estimation and the testing methods need a large number of observations to work well. Including all the tenders of the entire metropolitan area would increase the data set to about 300 auctions. That data is available but is not used. It can be argued that the suburban bus transit market is too different from the urban one for the data on these markets to be treated as identical. Also the scoring rules differed somewhat. The most important difference with the YTV data is that the bidders submitted many more combination bids than in the City of Helsinki auctions. Problems arising from scoring and combination bidding are discussed below.

The tendering process is as follows. First the Commercial Services Committee selects the service operators. Then the planning unit of Helsinki City Transport (HKL) and the City of Helsinki Supplies Department draw up and execute the competition process. HKL Bus Transport participates as one of the contractors. HKL planning unit decides routing, timetables, vehicle requirements and fleet schedules. HKL can change the amount of bus kilometers of a contract by a maximum of ten percent per year. The tenders are open to all contractors who have a licence to operate in the business. Also a financial analysis on the contractors' ability to fulfill the tender specifications is conducted.

In bids the operators state the unit costs of the service (cost per kilometer, per hour and per vehicle day), which the tendering authority uses to calculate the total cost of service provision which is the actual bid. HKL receives all ticket revenues. The contract period varies from three to six years being most often five years. The invitation to tender simultaneously covers many contracts. A single contract can cover one or more routes. The set of contracts that correspond to an invitation is called a tranche following Cantillon and Pesendorfer (2006a). Combination bidding within a tranche is allowed. (For more information on bus transit tenders in Helsinki region see for example YTV Transport Department (2001) or HKL (2006).)

The principle of awarding tenders is the best economic value that is calculated by a scoring rule based on the bids and vehicle quality. The lowest bid gets 86 points. Other bids' points are calculated with the formula: points for bid i = [lowest bid / bid i] * 86. A maximum of ten points are awarded for vehicle quality. The principle of giving points to different vehicle characteristics is clearly stated in the tender invitation. Two points are given for environment and quality certificates each. In the last three invitations in the data set the scoring rule had changed so that the bid was awarded a maximum of 87 points and vehicle quality 13 points, and the certificates were obligatory.

[Table 1. about here]

The data is described in table 1. The data consists of all the intra-city bus transit tenders held in the City of Helsinki in years 1997 - 2005. The data are collected by the author from the archives of the City of Helsinki Supplies Department (Saarelainen 2004). The data consists of 13 tranches, 64 contracts and 261 bids of which 14 were combination bids. The number of bidders varies from two to eight. Table 2. presents some descriptive statistics for each participating firm. Both the bid and the pseudo-cost results seem reasonable as firms with the highest percentage of wins have the lowest average bids and costs. The municipal bus company HKL participated in all auctions. Only five companies out of eleven were important players. They submitted 93 percent of bids and won all but one contract.

[Table 2. about here]

4.2 Applicability of the data

The data include information on bids, scores, contract characteristics and bidder characteristics. For the purposes of this study the information on bidders is not utilized with the exception of the location of garages. There are several problems with the application of HHS's tests to this data set. HHS assume that symmetric bidders bid in independent and homogenous auctions and that the number of bidders is exogenous and known to all bidders ex ante. The data is generated from combinatorial auction with asymmetric bidders and multiattribute bids. Other problems are small size of the data set, observed and unobserved contract heterogeneity and endogenous participation. The main problem arises from the different equilibrium behavior assumed in the estimation of the pseudo-costs and the actual auction considered here. Below some solutions are presented to alleviate the problems mentioned here. The discussion is useful in highlighting the possible limitations to applicability of the HHS tests. These limitations are mainly related to relying on nonparametric estimation of pseudo-costs.

As noted before, asymmetry makes the interpretation of the null as private costs impossible. We will see below that endogeneity and unobserved heterogeneity make rejecting this null harder even when it should be rejected. Thus when the null is not rejected it is difficult to be certain of the private costs environment even in the symmetric case. This is unfortunate because the alternative hypothesis of common costs incorporates many different models. Therefore if private costs are rejected, the econometrician needs to make assumptions about the exact form of the common costs environment to be able to conduct further study on the data. Also the policy considerations are more difficult if the null is rejected because for example the effects of more competition are unclear.

4.2.1 Multiattribute auction

The bids are multi-dimensional. As explained, bids consist of a monetary part and of a service quality part. They are transformed into one dimension in the following way. First the price of one point is calculated based on the scoring rule by dividing the winning bid with 86 (or 87 in the last three tranches). The same price is reached by dividing any bid with that bidder's price points. Then the quality points (points for vehicle quality and points for the certificates) that each bidder has received are multiplied by this price per point. That number is then subtracted from the bid. This method is based on the economic intuition that an optimally behaving winning bidder submits a bid where the cost of getting a quality point equals the expected cost of getting a monetary point.[2]

Asker and Cantillon (2006) have studied the properties of scoring auctions. They argued that making bidders' types one dimensional is sufficient for characterization of equilibrium when the scoring rule is quasi-linear and types are independently distributed. They also suggest that their paper provides a theoretical basis for the investigation of econometric identification of scoring auctions. However this problem is yet to be solved in the literature. Moreover, the scoring rule used in these bus transit auctions is not linear. Thus we need to adopt some arbitrary method like the one explained above. Cantillon and Pesendorfer (2006a) treat quality scores in their study as noise and use only the bids for identification. The minor importance of scores in the sense that they rarely change the order of bidders in my data

[2] This should hold for marginal points. If there are (dis)economies of scale in the production of quality this does not necessary hold for intra-marginal points. Berechman (1993, p. 121-123) argues that the bus industry in general operates under constant economies of scale with respect to fleet size.

would support this interpretation. The subtraction done above does not cause any problems in case the quality scores are just noise because then it is just a random downward scaling. The effect of this reduction vanishes when the HHS homogenization method below is applied. Moreover, if the scores are not random the above approach makes a correction that has at least the right direction.

4.2.2 Combination bids

In the City of Helsinki bus transit auctions the bidders can submit bids for packages of contracts within a single tranche. For example one bidder submitted in the first tranche (seven contracts) single bids for contracts 2,3,4,5,6 and 7 and a combination bid for the package consisting of contracts 6 and 7. LPV and CPV estimation do not allow for combination bids. The possibility of combination bidding changes equilibrium behavior. It also makes defining more competition much harder than in the case when only single bids are allowed. Cantillon and Pesendorfer (2006a) provide a method for identification of a combinatorial auction. Unfortunately this cannot be utilized here for two reasons. First, they assume the private costs paradigm and in this testing procedure we need the model to allow the possibility of affiliation and common costs. Second, the presence of combination bids is necessary for their identification method and these are rare in my data.

In this data there are 14 combination bids of which two were winning bids, winning a total of five contracts. Coincidentally 14 contracts out of 64 were included in these bids. Usually bidders included the same contracts in their combinations. For example in the first tranche, contract 6 was included in four combination bids submitted by three different firms, whereas only single bids were submitted for contract 1. A notable fact is that after the whole network had been procured once, bidders stopped submitting bids for packages even though the rules still allowed it. Thus it can be argued that combination bids do not play a large role in these auctions. There are three possible explanations for not submitting combination bids. First, there can be additional costs in calculating and submitting the combination bid. Second, there are negative cost synergies between contracts. Third, costs between the contracts are correlated. If either of the first two explanations is plausible, the auctions can be treated as independent and one could just omit the auctions with combinations. The third explanation is based on McAfee et al. (1989) who mention that it is always optimal for oligopolies to use bundling if the reservation values for products are independent. Identification with this correlation should be possible but should also be taken into account in the identification. At the moment this is left for further research. Now I just omit all the auctions that were included in any of the combination bids. I am left with 50 auctions. Then I assume that the bidders treated these 50 auctions as independent.

4.2.3 Observed contract heterogeneity

There are many important auction specific characteristics in the data. These shift the distribution of bidder costs and therefore have to be controlled for. HHS suggest two different ways to incorporate these observables. The first is the standard way to condition the estimation of pseudo-costs and the kernel smoothing on these observables. Because of the curse of dimensionality, this requires much larger data than available. The second alternative is to regress all the observed bids on the covariates and a set of dummy variables for each number of bidders. The sum of each residual and the corresponding intercept estimate provides an estimate of each homogenized bid. These estimates are then treated as bids in a sample of auctions of homogenous goods. In this study the bids are homogenized by regressing them against contract size in bus route hours, the share of articulated/bogie buses required for the contract, the share of rush hour traffic in the contract, the length of the contract in years and dummies for the number of bidders. This method requires the assumption that the cost structure is additively separable in contract characteristics. Please refer to HHS for details.

Table 3 shows that there is very large variation in the sizes of the bids and also in the sizes of the contracts. The largest bid is about 86 times larger that the smallest bid. The same is true for contract size. The other contract characteristic also vary a lot across contracts. Therefore the need for controlling for heterogeneity is obvious.

[Table 3. about here]

The results of the bid regression are in table 4. Contract characteristics explain about 97 % of the variation in the bids. Besides the dominant contract size also peak hour share and the share of articulated/bogie buses matter. The regression diagnostics (not shown) do not give rise to any large concerns. The effect of the number of bidders on bids is not statistically significant. When bids submitted in four or eight bidder auctions are compared to bids submitted in three bidder auctions, the effect is almost significant. There the increase in the number of bidders decreased the bids.

[Table 4. about here]

4.2.4 Unobserved contract heterogeneity

Observed characteristics explain most of the variation in bids. Still some important contract characteristics that affect the costs of all bidders in a similar way might not be included. If these are observed by the bidders the estimation of the pseudo-costs is biased. Krasnokutskaya (2004) has developed a method to take this into account. However her model is only identified under the private costs paradigm and is thus of no help here. Her results show that if the data generating process is actually a independent private costs one with unobserved heterogeneity, then using LPV or CPV estimation procedures leads to erroneous cost distributions. Wrongly estimated distributions have too low means and too high variances. It is not known whether the bias in the means is related to the number of bidders. Thus it is not clear how the lower means affect HHS tests. What we know is that higher variances may lead to situation where the null is not rejected even when it should be. My application turns out to be robust to this problem because the null is anyway rejected. However this is a serious consideration for other applications of HHS tests.

Also, there is a property in the data that alleviates this problem. Note that the entire traffic was tendered in the first seven tranches. Therefore the last six tranches hold, with the exception of some new routes and those auctions omitted due to combination bidding, exactly the same contracts as those in the first seven. Therefore there are ten identical contract pairs in the data. The only thing that is heterogenous within the pairs is that they were auctioned some years apart from each other. Of the ten pairs, five had a different number of bidders in them. The fact that some auctions are identical gives more credibility to the assumption of homogenous auctions.

4.2.5 Bidder asymmetry

In bus transit auctions bidder asymmetry arises mainly from different location of garages. The closer the routes are to the garages the less transfer kilometers the buses need to drive. Transfer means driving the empty bus from the garage to the start of the route at beginning of the shift of a particular bus and driving the bus back to the garage at the end of a shift. Bidders that have a garage near the contracted routes have an advantage. Asymmetry may also arise for other reasons such as capacity constraints. More free capacity increases incentives to bid more aggressively. Cantillon and Pesendorfer (2006a) argue that capacity effects may not be important for London bus route auctions, because firms have time to adjust their capacity between the auctions and the start of the contract traffic. Similar time lags are used in Helsinki. Also different collective labour agreements for public and private operators may affect the firms' relative ability to bid but these differences are marginal.

HHS's tests assume symmetric bidders. If bidders are asymmetric and their types are not observed one cannot distinguish between changes in the cost distributions resulting from different number of bidders and changes resulting from different sets of bidders. HHS suggest a way to extend their methods to detect common cost elements with asymmetric bidders. One would have to examine one bidder at a time and construct a sequence of sets of opponents faced by each bidder. This would require a much larger data set than available for this study. In this study I allow limited asymmetry by dividing the bidders into two groups based on the distance of their garage to the routes under contract and treating them as symmetric within groups. Asymmetry has to be taken into account both in the estimation of pseudo-costs and in testing as explained before.

The distances were calculated using an internet service (http://kartat.eniro.fi/) where it is possible to calculate road distances between street addresses in Helsinki. The gathering of results was automated using unix shell scripts. For contracts consisting of a single route the average distance from the nearest garage of the firm in question to the end points of the route was used. For contracts consisting of many routes the method varied based on the amount of traffic in particular routes. This somewhat arbitrary method is acceptable because it will not affect the selection to the groups. A median distance (11,5 kilometers) was used as a cut point in the grouping. Table 2. presents the distance descriptives for each bus company. Note that all the bids by HKL and AAS belong to the group with small distance and all the bids by OLA and the four companies with no garage in the area belong to the group with large distances.

4.2.6 Small data set, endogenous participation and pooling

Pooling means simply grouping some groups of observations. Pooling is used for two of the model specifications for two reasons. First, there are too few auctions with two and six bidders in the data for the use of subsampling. When the observations are pooled by treating the auctions with a small amount of bidders as belonging to one group and the auctions with a large amount of bidders as belonging to the other group, the entire data set can be used. Typically the results are also more robust to the choice of subsample size with larger samples. The second reason is that pooling alleviates the problem of possible endogenous participation as discussed below.

Pooling changes the distributions that are put to the tests in a following way. Let $n_p = 3$ denote the cut off number of bidders when auctions are pooled into two groups.

$$\text{(4.1)} \qquad \text{Pool 1: } \hat{F}_{v,n\leq n_p}(y) = \frac{1}{n_p-\underline{n}+1-\sum_{n=\underline{n}}^{n_p} 1\{T_n=0\}} \sum_{n=\underline{n}}^{n_p} \hat{F}_{v,n}(y)$$

$$\text{Pool 2: } \hat{F}_{v,n>n_p}(y) = \frac{1}{\bar{n}-n_p-\sum_{n=n_p+1}^{\bar{n}} 1\{T_n=0\}} \sum_{n=n_p+1}^{\bar{n}} \hat{F}_{v,n}(y),$$

where the indicator function $1\{\}$ gains value one if there are no auctions with a given number of bidders in the data.

Variation in the number of potential bidders, i.e. exogenous bidder variation, is most likely caused by observed bidder asymmetries in this market. A disadvantage in the location of the garage might make a bidder to perceive its probability of winning as too low to participate. Another possible reason are capacity constraints. Bidders observe each others' location and committed capacity. Therefore I assume that the number of observed bidders equals the number of potential bidders. These numbers can differ due to costs of calculating and submitting bids, a binding reserve price and unobserved contract heterogeneity. Explicit reserve prices are not used in these auctions but the tender invitation states the possibility of secret reserve price. I assume that the secret reserve price is not binding. This is plausible because the city owned company participates in all the auctions. Their bids act as a de facto reserve price from the perspective of the buyer. In theory the number of potential bidders is very large as the tender is open to any firm meeting certain standards. Therefore it could be difficult for both the bidders and even more so the researcher to know the number of potential bidders. However, to be able to participate competitively, a firm needs a garage in the area. Thus it is easy to detect possible new entrants. Moreover the number of participating players is quite small and the same players have been in the market for a long time. Thus it is reasonable to assume that most of the variation in the number of bidders is caused by observed exogenous factors. There is still a risk of putting some auctions to the wrong group when estimating the pseudo-costs. Firms could observe some potential bidders that in the end did not participate and that the researcher is unaware of. Pooling cannot be done at the stage of estimating the pseudo-costs but it helps at the testing stage. The chance of being put to the wrong group is reduced when pooling is used because the groups are larger. If bidder expected the number of participating bidders to be three but it turned out to be two, it is still in the correct pool.

HHS mention two ways in which endogenous participation may pose problems. First, if auctions with a large number of bidders tend to be those in which the contract is known by bidders to be particularly easy to operate, tests based on an assumption that variation in participation is exogenous can give misleading results. This would make observing the common costs paradigm more difficult In this application the tests suggest common costs so the results are robust to this problem. Second, nonparametric identification of the pseudo-costs depends on n being independent of any unobservables. HHS describe a structure under which both problems can be overcome using instrumental variables. Lacking the proper instrument it is assumed here that participation is exogenous. This is one potential cause of model misspecification that could drive the surprising result of common costs.

According to industry experts, small contracts that require from two to six regular 2-axel buses in which the share of weekend traffic relative to daytime weekday traffic is small are easier for all bidders, because the organization of traffic and fleet is easier. Also very high environmental requirements (EURO IV emission levels instead of the usual EURO III) make some contracts difficult for all bidders. Fortunately these are observed qualities. There are not many contracts of this type in the data. And as mentioned above, there is variation in the number of bidders within identical contracts suggesting that there must be at least some exogeneity in the process of participating in the bidding in the sense that the contract properties cannot explain all participation choices. Some exogeneity is created by mergers and acquisitions. Thus this problem is relatively small.

4.2.7 The equilibrium assumption

Another important consideration is that the pseudo-cost estimation is based on assuming Bayesian Nash equilibrium behavior. Some researchers find this strict rationality assumption implausible. As HHS admit, this is not an innocuous assumption, although they claim that FPSB auctions seem particularly well suited to this approach. Bajari and Hortacsu (2005) explore the reasonability of the equilibrium assumption and the structural approach by experimental data. Their results encourage the use of structural econometric tools. In this procurement setting the players are experienced companies bidding for large contracts. Thus the assumption of rational behavior is not far fetched. Kagel and Levin (1986) find out in their experiments that bidders learn equilibrium behavior only through experience. In my data the standardized variances of bids were much larger in the first tranche than later. This could imply learning. Due to the small size of the data set I still do not remove the first tranche from the data.

4.2.8 Collusion

Collusion would make testing impossible unless both all the bidders and the researcher knew the colluders and the type of collusion. It would change the equilibrium and make defining more competition difficult. In this case collusion is unlikely as it would yield very low rents. The most important bidder is a public company. Therefore it has no incentives to collude and its bids would bound the profits from colluding. The other fact that supports competitive bidding is the very low accounting profits that the bidders have in this market.

5 Empirical results

In this section I present the results of three different testing specifications. The first is the standard approach presented in HHS where there are symmetric bidders and pooling is not used. The second is the symmetric case with pooling and the third is the asymmetric case with pooling. I present the central results in the text. Robustness checks are presented in appendixes 4-6. Due to the small amount of observations with a given number of bidders, the actual data that is used in the analysis differs from one specification to another. The original data set is $T_n{}' = [6(n = 2), 19(n = 3), 11(n = 4), 11(n = 5), 2(n = 6), 1(n = 8)]$. I drop the lone eight bidder auction from all the specifications. I use only $T_n^{d\prime} = [19(n = 3), 11(n = 4), 11(n = 5)]$ in the standard approach with subsampling. When I use pooling the auctions with three or fewer bids are treated as belonging to one group and auctions with four to six bidders to the other group. I also use pooling in the asymmetric case. There I compare the 10 auctions with two near bidders and three far bidders (i.e. $n_0 = 2, n_1 = 3$) to the pool that consists of the 6 auctions with ($n_0 = 2, n_1 = 2$), the 5 auctions with ($n_0 = 1, n_1 = 2$) and the 8 auctions with ($n_0 = 2, n_1 = 1$). This guarantees that there is no ambiguity in comparing the distributions, because the group with a large number of bidders has at least an equal number of bidders of one type and more of the other type than all the auctions in the group with the smaller number of bidders. The means test is conducted with all these testing specifications.

5.1 Testing specification 1: The standard approach

5.1.1 Auctions with two to six bidders

In this model all the distributions for different numbers of bidders are treated separately. Asymmetry is not taken into account. It is not clear from figure 1. whether one should suspect common costs. Based on the picture alone the stochastic dominance is not clear as the lines cross. On the other hand, for the majority of observations $\hat{F}_{v,n}(v) > \hat{F}_{v,n+1}(v) > ... > \hat{F}_{v,\bar{n}}(v)$ holds.

[Figure 1. about here]

For this model I use only the means test because there a too few two and six bidder auctions for the use of the subsampling procedure. The test statistics and the p-values for different quantiles are presented in table 5. The null hypothesis of private costs is not rejected when the quantile is chosen to be 5 %. When the trimmed amount is increased the null is rejected. This is due to decreasing bootstrapped standard deviations.

It is very important to note that there are qualitative changes in the inference that depends on the choice of the amount trimmed. HHS assume that the choice of quantile is arbitrary as it does not matter asymptotically. A data driven method for choosing the quantile should be formulated. One possibility would be to look at simulations for a reasonable quantile to trim using the same sample size and bandwidth selection rule as in estimating the pseudo-costs. Monte Carlo simulations would reveal whether we could find some rules of thumb for choosing the quantile.

[Table 5. about here]

5.1.2 Auctions with three, four or five bidders.

To be able to use the sup-norm test I conduct the standard testing procedure also with a smaller data set. I use only auctions to which three to five bidders have participated. Again there is non-conclusive evidence against the null in figure 2. For most of the values of v $\hat{F}_{v,n=3}(v) < \hat{F}_{v,n=4}(v)$ implying that the null cannot be rejected. Then again $\hat{F}_{v,n=5}(v)$ gains the smallest values of the three distributions through most of the range of v, implying rejection.

[Figure 2. about here]

Compared to the previous exercise, the means test for the sample with only 3-5 bidder auctions gives larger p-values. These are mostly due to dropping the six bidder auctions. Again the means test is not robust to the choice of the quantile. The rejection of the null depends on the trimming choice. The sup-norm test clearly rejects the null and is quite robust as can be seen from appendix 3.

[Table 6. about here]

Many choices concerning subsampling that are arbitrary can affect the results. These choices are the subsample size, the number of subsamples taken and which repetition of the test to use. Most notably the choice of the subsample size can change the results. In theory, the subsample size should be far away from both 1 and T. Linton et al (2005) suggest computing a plot of p-values against subsample sizes for a range of subsample sizes. If the p-value is insensitive to subsample size within a "reasonable" range, then inferences are likely to be robust. Regarding the number of subsamples taken there is a trade-off between the computer time and the robustness of the result. We can be pretty sure of getting a correct and robust result with 5000 draws but not necessarily with 50 draws. The p-value also changes from one push of the button to the next. Therefore one should also check whether the p-value is robust to repeating the test with the chosen subsample size and the number of draws. Appendixes 3-5 present these robustness checks for the different specifications. It is important to be aware of these problems with robustness and to discuss the choices that are made. Monte Carlo simulations could reveal important information on how to make these choices.

5.2 Testing specification 2: Pooling with symmetry

Now I pool the auctions into two groups. The first pool includes the auctions with two or three bidders, and the second pool includes the auctions with four, five and six bidders. The distributions for the two pools are presented in figure 3. It gives no clear evidence concerning the information paradigm as the lines cross. The test statistics and the p-values are in table 7. Pooling changes the results of the means test significantly. This time it does not reject the null. The sup-norm test gives very similar values as before and rejects the null at the 5 % significance level. Both tests seem to be very robust to the arbitrary choices.

[Figure 3. about here]

[Table 7. about here]

The results from the sup-norm tests for all the symmetric specifications clearly indicate that the bidders operate in a common costs environment. The means tests give ambiguous results. HHS indicates the sup-norm test as their preferred test. I conclude that the symmetric results show evidence of common costs. This leads to clear answers to the policy questions of this study. The merger of the two bus companies might have no effects, or even a decreasing effect, on the procurement costs due to decreasing competition. For sure it has a lesser increasing effect than feared. Thus from the perspective of competition this merger cannot be criticized. The reasons for the merger remain valid. These include getting more distance between the planner (HKL planning unit) and one bidder (HKL bus transit), synergies and getting rid of having two companies with the same owner in these auctions. The other policy activity currently under consideration, that is building of city owned garages and renting them to new entrants to induce more competition, should not necessarily be undertaken. It is not clear that new competition brings any benefits to the city in the form of savings in procurement costs. Therefore the city might be better off not spending any resources to induce new competition.

5.3 Testing specification 3: Pooling with asymmetry

To control for asymmetry the observations are grouped according to the distance from a given bidder's nearest garage to the route under contract. Both tests are then conducted separately for the different groups. Now the pseudo-costs are estimated using the asymmetric set up. I omitted one auction with one near bidder and two far bidders from the sup-norm test from the data because it had one unusually small outlier value for the pseudo-cost. Table 8 presents the pooling rule I used in the testing.

[Table 8. about here]

Figure 4 shows the distributions for the small distance group. It clearly implies common costs because the distribution for the pool with a small number of bidders gets larger or about equal values throughout the entire range of observed values. This is confirmed by the test results in table 9. Both tests reject the null hypothesis of private costs. Both tests work well from the robustness point of view for both of the distance groups as can be seen from appendix 5.

[Figure 4. about here]

[Table 9. about here]

Figure 5 presents the distributions for the two pools for the bidders with garages located far from the routes. It shows very clearly that there is no need even to test the difference in distributions for this group as there is no observable difference in them. It is still interesting to compare the p-values with other testing specifications and thus the tests are nonetheless conducted. The results in table 10 confirm the expectations. Neither of the tests rejects the null of equal distributions. The means test gives higher p-values.

[Figure 5. about here]

[Table 10. about here]

Asymmetric results mean that the bus companies operating from garages far from the routes could operate either in the private or the common cost environment and bus companies that have garages near the routes operate in the common cost environment. We also know that even if also far operators had common costs, the common cost elements are more important to the near bidders. Thus if there are any procurement cost reducing the effects of competition, they are larger for the far operators. All the conclusions made for the symmetric case remain valid with asymmetry taken into account. Moreover it is now possible to make additional conclusions on the two policy questions.

First, the two merged companies were more often near than far bidders. Thus the merger reduced the number of bidders to whom the common cost elements were more important. This actually makes the policy of inducing more competition more plausible because now the role of common elements in the market is diminished. Second, if the city chooses to pursue the proposed garage policy, and they have to choose between two locations from which the overall operation costs are equal, they should choose the garage which is more often located further away from the routes. Then at least the new entrants react to the increased competition themselves. For the behavior of the incumbents the location does not matter because assuming similar overall operating costs, incumbents do not perceive the new bidders or locations as strong or weak. The only difference is along the common or private costs dimension.

6 Conclusions

Due to the small number of observations in the data relative to the requirements of the econometric methods and the arbitrary treatment applied to the data to take combination bidding and the scoring rule into account, the results of this study should treated be with caution. The means test was not robust to the choice of the quantile trimmed in the symmetric case in a sense that the p-values varied a lot with it and there were also qualitative changes. HHS assume that the choice of the quantile is arbitrary as it does not matter asymptotically. For the sup-norm test robustness checks should be conducted with regards to the subsample size, the number of subsamples taken and for different repetitions of the test. The results for the symmetric testing approach show that there seem to be important common cost elements in this market. The asymmetric approach gave very clear and robust results. It seems that the bus companies that have garages close to the contracted routes operate in an common cost environment.

Because the equal distributions hypothesis was not rejected for bus companies that have garages far from these routes, the information environment is not known for them. Common costs elements are more important to operators with garages near the routes. Common costs can arise from common future uncertainty and private costs from individual efficiency differences. Next I provide explanations for this result.

The garages that are most often located near the routes are also closer to the city center where the costs of the land needed for the garage is higher. Land rents or opportunity costs of land are subject to future uncertainty. Unlike for example gasoline price changes, changes in the land rents are not covered by the contract terms. For example an operator with obligations to operate routes requiring a total of 50 buses for the next five years could face significant changes in production costs if the land rents increase. Another factor is that being located near the route makes the incentives to be efficient in organizing the empty transfer traffic lower. This makes the private cost component less important. When garages are far the uncertainty concerning the rents are smaller and the need to be more efficient is higher.

Another element that could be driving common costs are outside options. Bidders also participate in auctions for the metropolitan transit. When they commit their garages to traffic in one contract they reduce their changes to participate and win in following auctions. Uncertainty about winning future contracts is common. There could also be common uncertainty about the results of future negotiations with the labour unions. There could be strikes or some new frictions that are not covered by the contract. The industry also suffers from undersupply of driver labour (Helsingin Sanomat 29.1.2007 and 12.2.2007). Therefore it is not clear whether the bidders get enough drivers if they win contracts. This also creates common uncertainty.

Valkama and Finkkilä (2003) study the economic effects of starting tendering of bus services in the Helsinki Metropolitan Area. They analyze firm accounting data from 1998 to 2001. They found that tendering reduced the firms' profits dramatically. For a while firms made positive total profits because of the old negotiated contracts but the tendered traffic induced losses. In year 2001 the firms operated with losses. For example operating margin, net income and return on capital were negative. In 2003 the firms on average managed to break even and they have made small profits after that (see for example (HKL 2006)). This time series can be evidence of aggressive competition for market shares in the beginning. But it also could be evidence for common costs. It took time for firms to learn to take the winner's curse effect into account sufficiently enough to avoid losses.

The symmetric result leads to clear answers to the policy questions posed in this study. A decrease in competition caused by the merger of the two bus companies could have no effects or even an reducing effect on the procurement costs. For sure it has a lesser increasing effect than feared. Thus from the perspective of competition this merger decision cannot be criticized. The reasons for the merger remain valid. These include getting more distance between the planner (HKL planning unit) and one bidder (HKL bus transit department), synergies and getting rid of having two companies with the same owner in these auctions. The other policy activity currently under consideration, that is building of city owned garages and renting them to new entrants to induce more competition, should not be undertaken. It is not clear that new competition brings any benefits. Therefore the city should not spend any resources to induce new competition. Asymmetric results make it possible to make additional conclusions on these two policy questions. First, the two merged companies were more often near than far bidders. Thus the merger reduced the number of bidders to whom the common cost elements were more important. This actually makes the policy of inducing more competition more plausible because now the role of common elements in the market is diminished. Second, if the city chooses to pursue the proposed garage policy, and they have to choose between two locations from which the overall operation costs are equal, they should choose the garage which is more often located further away. Then at least the new entrants react to the increased competition themselves. For the behavior of the incumbents the location does not matter because assuming similar overall operating costs, incumbents do not perceive the new bidders or locations as strong or weak in general.

Table 1. The bus transit tenders included in the data set.

Tranche	# auctions	Size	# S bids	# C bids	min - max
1/I (97-98)	7	4.9	34	5	3-8
1/IA (98)	1	0.02	4	0	4
1/II (98)	7	7.5	23	1	2-6
1/III (98-99)	1	0.4	8	0	8
1/IV (99)	9	7.7	33	4	2-6
1/V (00)	8	7.5	25	3	2-5
1/VI (01)	5	5.4	18	1	2-4
2/I (01)	3	2.8	11	0	3-5
2/II (02)	5	5.6	21	0	3-5
2/III (03)	5	0.7	20	0	3-5
2/IV (03)	4	3.6	18	0	3-5
2/V (05)	8	7.9	27	0	3-4
2/VI (05)	1	0.1	5	0	5
Total	64	54	247	14	2-8

"Tranche" refers to the set of contracts that correspond to an single invitation to tender. The year the auction was held is in parentheses. 2/III means the third tranche of the second round of tendering of the entire traffic. "# auctions" refers to the number of auctions in a given tranche. Size gives the total size of all the contracts in a given tranche in millions of bus kilometers in a year. "# S bids" means the total number of single bids in a given tranche. "# C bids" means the total number of combination bids in a given tranche. "min - max" refers to the spread of the number of bidders per auction in a given tranche.

Table 2. Descriptive statistics on the distances from garages to routes, the homogenized bids and the estimated pseudo-costs for each participating firm.

Firm	HKL	CX	STA	CR	PKL	OLA
# bids	50	46	26	37	16	4
# winning bids	19	5	8	12	5	0
Distance from the nearest carage to the contracted route(s) (km)						
mean	5,9	13,7	7,3	13,0	18,0	14,2
std. deviation	1,9	2,3	3,3	2,6	2,3	1,2
min	2,9	6,9	3,3	8,5	13,5	12,6
max	10,5	17,7	20,0	20,1	20,9	15,1
Homogenized quality adjusted bids						
mean	99	187	120	110	56	529
std. deviation	223	187	196	177	101	570
min	-417	-86	-243	-226	-101	1
max	1234	751	645	699	239	1256
Pseudo-costs assuming symmetry						
# missing values*	1	0	0	0	0	2
mean	-160	55	-16	-69	-78	46
std. deviation	456	260	329	302	211	64
min	-2588	-431	-818	-747	-460	1
max	335	751	645	699	239	91
Pseudo-costs assuming asymmetry						
# missing values*	22	23	10	16	7	3
mean	-413	29	-64	-72	-31	159
std. deviation	1498	178	312	198	193	NA
min	-7892	-345	-681	-365	-259	159
max	231	297	645	279	239	159
Firm	**LLR**	**AAS**	**LSL**	**ESL**	**AAD**	**Total**
# bids	2	2	2	2	1	188
# winning bids	0	0	0	1	0	50
Distance from the nearest carage to the contracted route(s) (km)						
mean	NA	6,2	NA	NA	NA	11,3
std. deviation	NA	1,0	NA	NA	NA	5,4
min	NA	5,5	NA	NA	NA	2,9
max	NA	7,0	NA	NA	NA	20,9
Homogenized quality adjusted bids						
mean	579	97	205	115	301	139
std. deviation	77	20	283	143	NA	216
min	524	83	5	14	301	-417
max	633	111	405	216	301	1256
Pseudo-costs assuming symmetry						
# missing values*	1	0	1	0	0	5
mean	524	85	5	26	258	-46
std. deviation	NA	3	NA	269	NA	343
min	524	83	5	-164	258	-2588
max	524	87	5	216	258	751
Pseudo-costs assuming asymmetry						
# missing values*	2	2	2	0	0	87
mean	NA	NA	NA	19	301	-131
std. deviation	NA	NA	NA	280	NA	820
min	NA	NA	NA	-179	301	-7892
max	NA	NA	NA	216	301	645

HKL = Helsingin Kaupungin Bussiliikenne. CX = Connex Oy. STA = Suomen Turistiauto. CR = Concordia Oy. PKL = Pohjolan Kaupunkiliikenne. OLA = Oy Liikenne Ab. LLR = Linjaliikenne Randell. AAS = Askaisten Auto. LSL = LS-Liikennelinjat Oy. ESL = Etelä-Suomen Linjaliikenne. AAD = Auto Andersson Oy. *The five missing values for the symmetric case are due to zeros in the denominator

$\hat{g}^*_{B_1,n}(b;b)$. The statistical program gives zero for some values of this kernel density estimator, because there are very few observations spread over a large area. Some of the missing values for the asymmetric case are also caused by this. Most of them however arise from the fact that I estimate the asymmetric case only for 34 auctions whereas the symmetric case is estimated for 50 auctions.

Table 3. Descriptive statistics of the variables used in the homogenizing regression.

Variable	Mean	Std. Dev.	Min	Max
Quality adjusted bids (mil. 2001 €)	1.75	1.16	0.06	5.19
Route hours per year (thousands)	42.5	27.8	1.3	111.2
Share of peak hour traffic	0.39	0.23	0	1
Share of articulated/bogie buses	0.23	0.36	0	1
Contract length (years)	4.78	0.69	3	6

In addition to these variables there are also dummies for the number of bidders. There are 6 two bidder auctions, 19 three bidder auctions, 11 four bidder auctions, 11 five bidder auctions, 2 six bidder auctions and 1 eight bidder auction in the data set used for this regression.

Table 4: The OLS results of explaining bids with contract characteristics. The bids are in units of 1000 euros.

Coefficient	Estimate	Std. Error	P-value
Intercept	190	146	0.19
Route hours	*0.04*	*0.00073*	*<2e-16*
Peak hour share	*196*	*84.9*	*0.022*
Artic./Bogie buses share	*179*	*51*	*0.00057*
Contract length	-47.2	29.6	0.11
n=2	-42.4	69	0.54
n=3		Reference group	
n=4	-88.2	45.5	0.054
n=5	-53.6	42.3	0.21
n=6	*-88.6*	*71.9*	*0.21*
n=7		No observations	
n=8	-166	99.4	0.096
Adjusted R^2		0.965	
F-statistic		576 (9 and 178 df)	

Figure 1. The empirical cumulative distributions of pseudo-costs for different numbers of bidders.

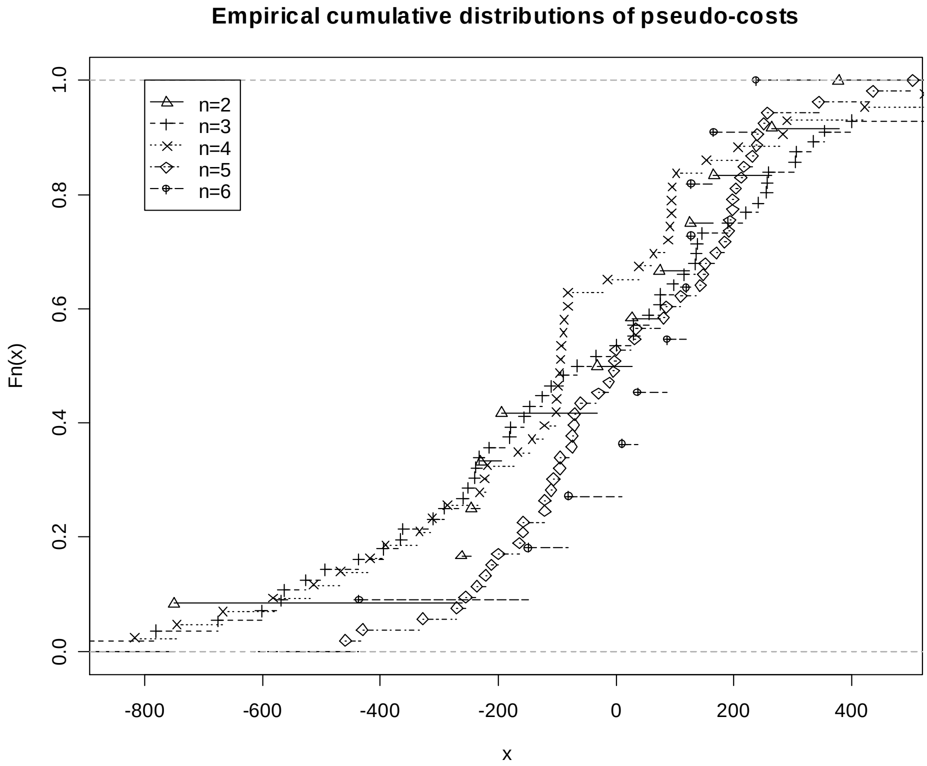


Figure 1. draws $\hat{F}_{v,n}(y) = \frac{1}{T_n \times n} \sum_{t=1}^{T} \sum_{i=1}^{n} 1\{\hat{v}_{it} < y, n_t = n\}$ separately for each $n = 2, ..., 6$.

Table 5. Quantile trimmed means, bootstrapped standard deviations, the test statistics and the p-values from the means test for the full sample.

n	T	qtm	sd	X^2-bar	p-value
		quantile=5%			
2	6	-26	61		
3	19	-34	46		
4	11	-76	44	6.14	0.07084
5	11	29	39		
6	2	56	36		
		quantile=10%			
2	6	-26	34		
3	19	-35	36		
4	11	-67	28	17.32	0.00037
5	11	24	29		
6	2	49	13		
		quantile=20%			
2	6	-19	27		
3	19	-21	22		
4	11	-47	18	38.01	1.9E-08
5	11	20	18		
6	2	42	0.9		

"n" refers to the number of bidders and "T" to the number of auctions with given number of bidders. "qtm" gives the quantile trimmed mean and "sd" its bootstrapped standard deviation. "χ^2-bar" is the test statistic for the given quantile.

Figure 2. The empirical cumulative distributions of pseudo-costs for three, four and five bidders.

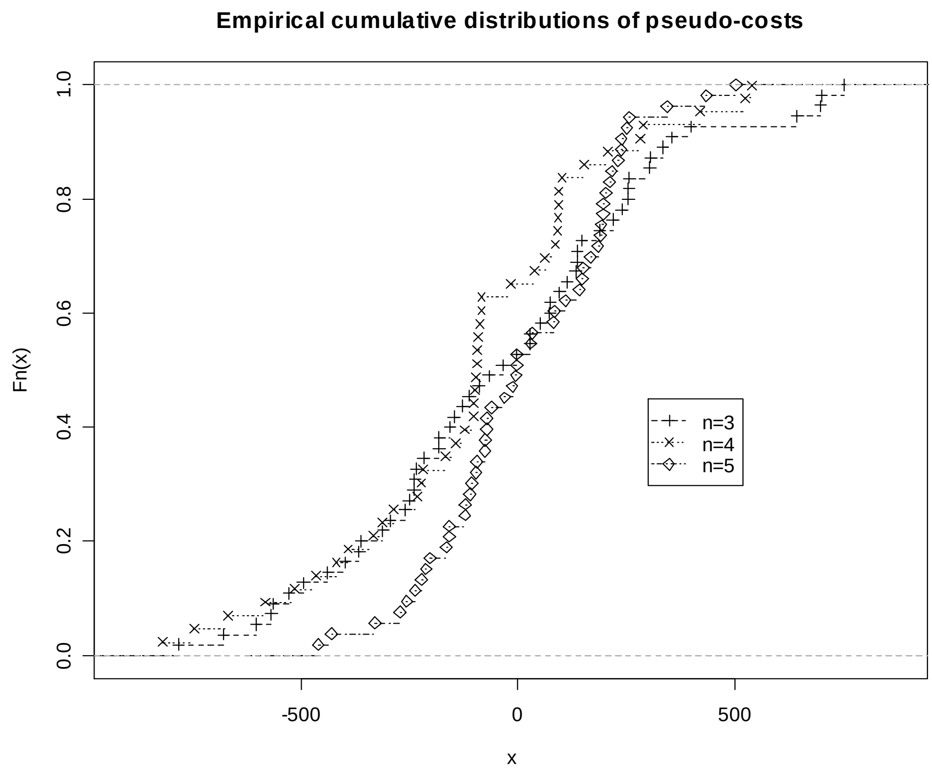
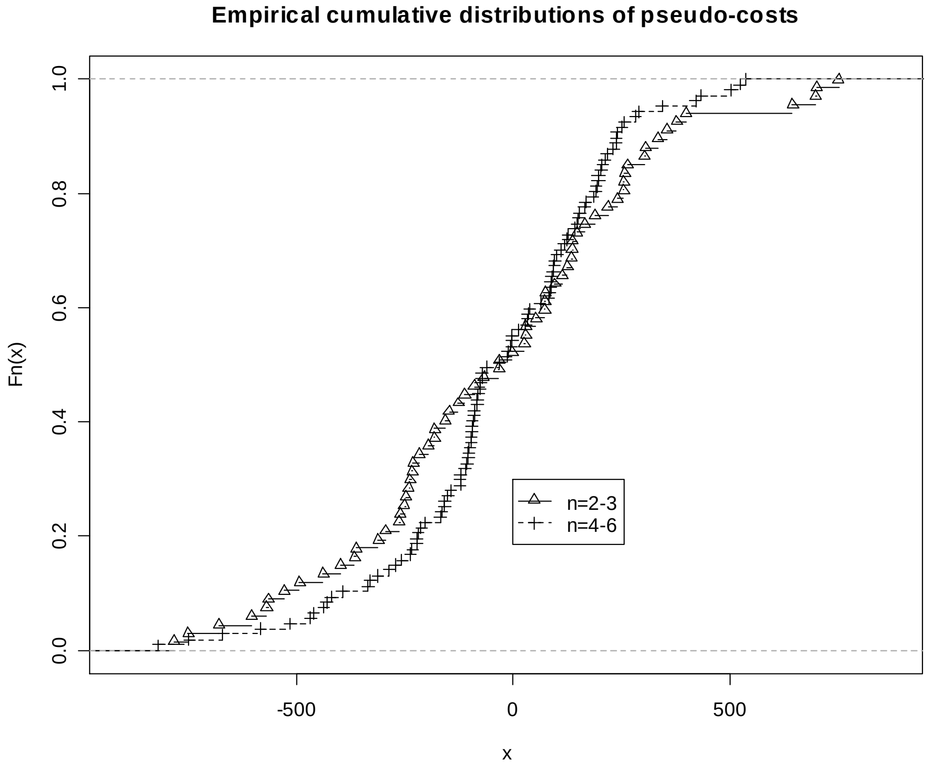


Table 6. The test statistics and the p-values from the means and the sup-norm tests for auctions with 3-5 bidders.

	Means test		
	Quantile 5 %	**Quantile 10 %**	**Quantile 20 %**
X^2-bar	2.87	4.71	6.38
p-value	0.224	0.096	0.043
	Sup-norm test (subsample sizes 9 and 6)		
p-value		0.006	

Figure 3. The empirical cumulative distributions of pseudo-costs for different numbers of bidders, with two to three and four to six bidder auctions pooled.



The distribution for the pool 1 is $\hat{F}_{v,n\leq 3}(y) = \frac{1}{2}\sum_{n=2}^{3}\hat{F}_{v,n}(y)$ and for the pool 2 $\hat{F}_{v,n>3}(y) = \frac{1}{3}\sum_{n=4}^{6}\hat{F}_{v,n}(y)$.

Table 7. The test statistics and the p-values from the means and the sup-norm tests for the pooled groups.

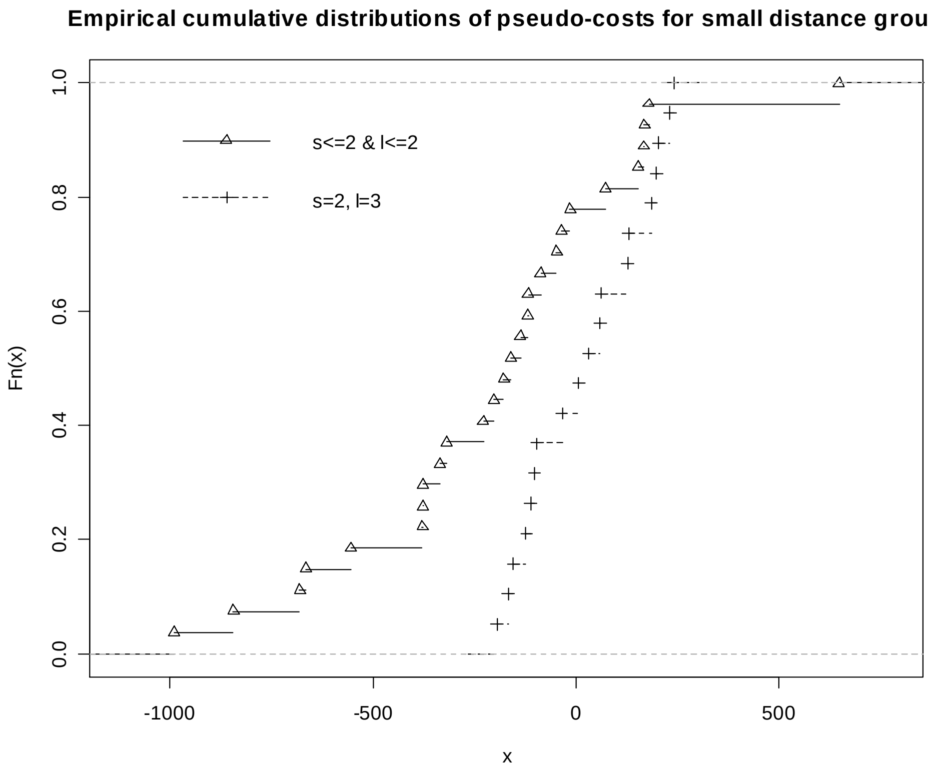
Means test			
	Quantile 5 %	**Quantile 10 %**	**Quantile 20 %**
X^2-bar	0.22	0.40	0.45
p-value	0.76	0.68	0.64
Sup-norm test (both subsample sizes 9)			
p-value		0.014	

Table 8. Pooling rule and number of auctions for each distribution under comparison.

pool	d	s	l	n	T	# bids	T_{pool}
1	large	2	3	5	10	30	10
2	large	2	2	4	6	12	
2	large	2	1	3	5	5	18
2	large	1	2	3	7	14	
3	small	2	3	5	10	20	10
4	small	2	2	4	6	12	
4	small	2	1	3	5	10	18
4	small	1	2	3	7	7	

The auctions are pooled into four groups. For example pool one consists of bidders with large (over 11.5 km) distance ("d") from garage to a given route in auctions with number of short distance bidders "s" two and number of large distance bidders "l" three. "n" denotes the total number of bidders. "T" denotes the number of auctions of given type and "# bids" the number of bids that goes to a given pool from a given auction. "T_{pool}" denotes the number of auctions in a given pool.

Figure 4. The empirical cumulative distributions of pseudo-costs for bidders belonging to the small distance group G0. 10 auctions with $(n_0 = 2, n_1 = 3)$ are compared to pool that consists of 6 auctions with $(n_0 = 2, n_1 = 2)$, 5 auctions with $(n_0 = 1, n_1 = 2)$ and 8 auctions with $(n_0 = 2, n_1 = 1)$.



Distribution for pool 1 is $\hat{F}_{v,(n_0,n_1)=(2,3),d\leq 11,5}(y)$ and for pool 2

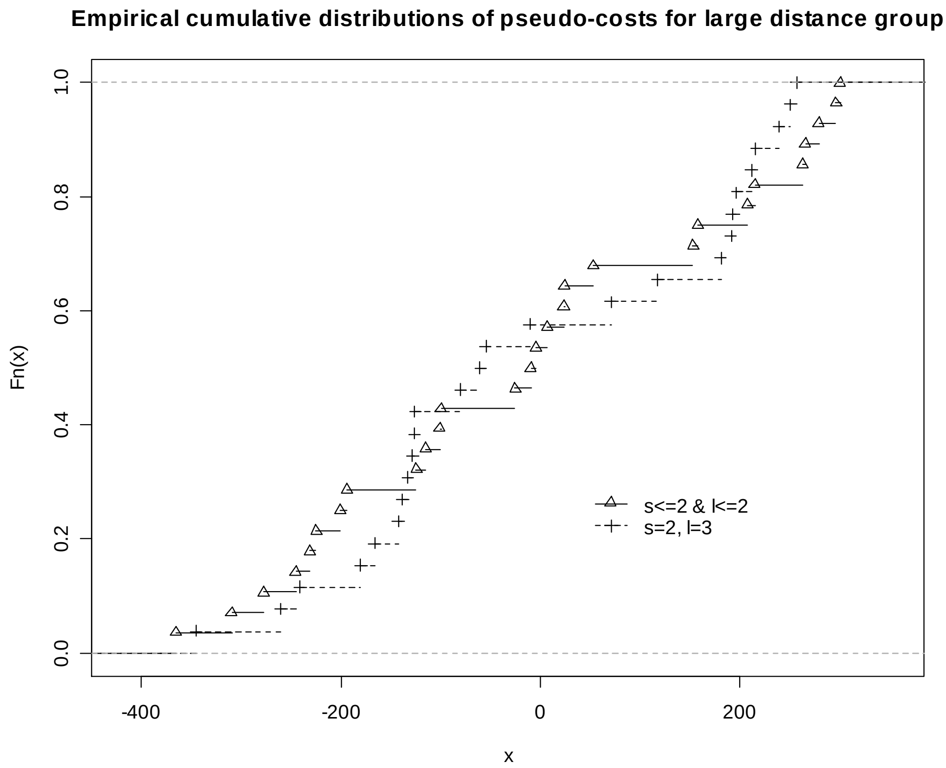
$\frac{1}{3}[\hat{F}_{v,(n_0,n_1)=(2,2),d\leq 11,5}(y) + \hat{F}_{v,(n_0,n_1)=(2,1),d\leq 11,5}(y) + \hat{F}_{v,(n_0,n_1)=(1,2),d\leq 11,5}(y)]$.

$\hat{F}_{v,(n_0,n_1)=(s,l),d\leq 11,5}(y) = \frac{1}{T_{(n_0,n_1)=(s,l)} \times n_1} \sum_{t=1}^{T_{(n_0,n_1)=(s,l)}} \sum_{i=1}^{n_1} 1\{\hat{v}_{1it} < y, (n_0, n_1) = (s, l)\}$.

Table 9. The test statistics and the p-values from the means and the sup-norm tests for the group with garages located near the routes.

Means test			
	Quantile 5 %	Quantile 10 %	Quantile 20 %
X^2-bar	6.04	8.67	14.77
p-value	0.032	0.008	0.0004
Sup-norm test (subsample sizes 9 and 5)			
p-value		0.002	

Figure 5. The empirical cumulative distributions of pseudo-costs for bidders belonging to the large distance group G1. 10 auctions with $(n_0 = 2, n_1 = 3)$ are compared to a pool that consists of 6 auctions with $(n_0 = 2, n_1 = 2)$, 5 auctions with $(n_0 = 1, n_1 = 2)$ and 8 auctions with $(n_0 = 2, n_1 = 1)$.



Distribution for pool 3 is $\hat{F}_{v,(n_0,n_1)=(2,3),d>11,5}(y)$ and for pool 4

$\frac{1}{3}[\hat{F}_{v,(n_0,n_1)=(2,2),d>11,5}(y) + \hat{F}_{v,(n_0,n_1)=(2,1),d>11,5}(y) + \hat{F}_{v,(n_0,n_1)=(1,2),d>11,5}(y)]$.

$$\hat{F}_{v,(n_0,n_1)=(s,l),d>11,5}(y) = \frac{1}{T_{(n_0,n_1)=(s,l)} \times n_0} \sum_{t=1}^{T_{(n_0,n_1)=(s,l)}} \sum_{i=1}^{n_0} 1\{\hat{v}_{0it} < y, (n_0, n_1) = (s, l)\}.$$

Table 10. The test statistics and the p-values from the means and the sup-norm tests for the group with garages located far from the routes.

Means test			
	Quantile 5 %	**Quantile 10 %**	**Quantile 20 %**
X^2-bar	0.09	0.25	0.00
p-value	0.86	0.75	1.00
Sup-norm test (subsample sizes 9 and 5)			
p-value		0.53	

7 References

Asker, J. - Cantillon, E. (2006): Properties of Scoring Auctions. Working Paper.

Athey, S. - Haile, P. (2005): Nonparametric Approaches to Auctions. Forthcoming in Heckman, J. - Leamer, E.,
eds. Handbook of Econometrics, Vol 6.

Ayer, M. - Brunk, D. - Ewing, G. - Reid, W. - Silverman E. (1955): An Empirical Distribution Function for Sampling with Incomplete Information. Annals of Mathematical Statistics 26, 641-647.

Bajari, P. - Hortacsu, A. (2005): Are Structural Estimates of Auction Models Reasonable? Evidence from Experimental Data. Journal of Political Economy 113, 703–741.

Bartholomew, D. (1959): A Test of Homogeneity for Ordered Alternatives. Biometrika 46, 36-48.

Berechman, J. (1993): Public Transit Economics and Deregulation Policy. Studies in Regional Science and Urban Economics 23. North-Holland, Amsterdam.

Campo, S. - Perrigne, I. - Vuong, Q. (2003): Asymmetry in First-Price Auctions with Affiliated Private Values. Journal of Applied Econometrics 18, 179-207.

Cantillon, E. - Pesendorfer, M. (2006a): Combination Bidding in Multi-Unit Auctions. Mimeo, London School of Economics.

Cantillon, E. - Pesendorfer, M. (2006b): Auctioning Bus Routes: The London Experience. Ch. 23 in Cramton, Shoham and Steinberg (eds), Combinatorial Auctions. MIT Press.

Chernozhukov, V. - Fernandez-Val, I. (2005): Inference on Quantile Regression Process. Sankhya, The Indian Journal of Statistics 67, 253-276.

Donald, S. - Paarsch, H. (1993): Piecewise Maximum Likelihood Estimation in Empirical Models of Auctions. International Economic Review 34, 121-148.

Donald, S. - Paarsch, H. (1996): Identification, Estimation, and Testing in Parametric Empirical Models of Auctions Within the Independent Private Values Paradigm. Econometric Theory 12, 517-567.

Elaykime, B. - Laffont J-J. - Loisel, P. - Vuong, Q. (1994): First-Price Sealed-Bid Auctions with Secret Reservation Prices. Annales d'Economie et de Statisque 34, 115-141.

Gilley, O. - Karels, G. (1981): The Competitive Effect in Bonus Bidding: New Evidence. The Bell Journal of Economics 12, 637-648.

Guerre, E. - Perrigne, I. - Vuong, Q. (2000): Optimal Nonparametric Estimation of First-Price Auctions. Econometrica 68, 525-574.

Haatainen, J. (2003): Bussiliikenteen kilpailuttamisen henkilöstövaikutukset. Tutkimuskohteena pääkaupunkiseudun bussiliikenne. Tampereen yliopisto. Hallintotieteen laitos. Sarja A/31.

Haile, P. A. - Hong, H. - Shum, M. (2003): Nonparametric Tests for Common Values in First-Price Sealed-Bid Auctions. Working Paper.

Helsingin Sanomat (29.1.2007). Helsingin Bussiliikenteeltä jäi ajamatta 25 000 kilometrin verran vuoroja. Juha Salonen.

Helsingin Sanomat (12.2.2007). Helsinki alkaa kouluttaa nuoria bussinkuljettajiksi. Matti Huhta.

Hendricks, K. - Paarsch, H. J. (1995): A Survey of Recent Empirical Work Concerning Auctions. Canadian Journal of Economics 28, 403-426.

Hendricks, K. - Pinkse J. - Porter R. (2003): Empirical Implications of Equilibrium Bidding in First-Price, Symmetric, Common-Value Auctions. Review of Economic Studies 70, 115-146.

Hendricks, K. - Porter, R. (2005): Lectures on Auctions: An Empirical Perspective. Handbook of industrial Organization, Vol. III. North-Holland.

HKL (2006): Helsinki City Transport Annual Report 2005.

Hong, H. - Shum, M. (2002): Increasing Competition and the Winner's Curse: Evidence from Procurement. Review of Economic Studies 69, 871-898.

Kagel, J. (1995): Auctions: A Survey of Experimental Research. In The Handbook of Experimental Economics by Kagel and Roth (editors). Princeton University Press. 501-585.

Kagel, J. - Levin, D. (1986): The Winner's Curse and Public Information in Common Value Auctions. American Economic Review 76, 894-920.

Krasnokutskaya, E. (2004): Identification and Estimation in Highway Procurement Auctions Under Unobservable Auction Heterogeneity. Working Paper. University of Pennsylvania.

Krishna, V. (2002): Auction Theory. Elsevier Science.

Laffont, J-J. (1997): Game Theory and Empirical Economics: The Case of Auction Data. European Economic Review 41, 1-35.

Laffont, J-J. - Vuong, Q. (1996): Structural Analysis of Auction Data. American Economic Review 86, 414-420.

Laffont, J-J. - Ossard, H. - Vuong, Q (1995): Econometrics of First-Price Auctions. Econometrica 63, 953-980.

Li, T - Perrigne I. - Vuong, Q. (2002): Structural Estimation of the Affiliated Private Value Auction Model. RAND Journal of Economics 33, 171-193.

Linton, O. - Maasoumi, E. - Whang, Y-J. (2005): Consistent Testing for Stochastic Dominance Under General Sampling Schemes. Review of Economic Studies 72, 735-765.

McAfee, P. - McMillan, J. - Whinston, M. (1989): Multiproduct Monopoly, Commodity Bundling, and Correlation of Values. The Quarterly Journal of Economics 104. 371-383.

Milgrom, P. (2004): Putting Auction Theory to Work. Cambridge University Press.

Milgrom, P. R. - Weber, R. J. (1982): A Theory of Auctions and Competitive Bidding. Econometrica 50, 1089-1122.

Paarsch, H. J. (1992): Deciding Between the Common and Private Value Paradigms in Empirical Models of Auctions. Journal of Econometrics 51, 191-215.

Paarsch, H. J. - Hong, H. (2006): An Introduction to the Structural Econometrics of Auction Data. MIT Press.

Perrigne, I. - Vuong, Q (1999): Structural Econometrics of First-price Auctions: A Survey of Methods. Canadian Journal of Agricultural Economics 47, 203-223.

Pinkse, J. - Tan, G. (2005): The Affiliation Effect in First-Price Auctions. Econometrica 73, 263-277.

Saarelainen, K. (2002): Esipuhe tutkimukseen: Tukiainen, Janne (2002): Helsingin sisäisen bussiliikenteen kilpailuttamisen empiirinen tarkastelu huutokauppateorian näkökulmasta. Pro Gradu – opinnäytetyö. Helsingin yliopisto, Suomi.

Saarelainen, K. (2004): Helsingin kaupungin bussiliikenteen kilpailutuksen tarjouspyyntö- ja tarjousvertailuaineisto vuosilta 1997-2004.

Sareen, S. (1999): Posterior Odds Comparison of a Symmetric Low-Price, Sealed-Bid Auction Within the Common Value and Independent Private Value Paradigms. Journal of Applied Econometrics, 41, 205-235.

Shneyerov, A. (2003): An Empirical Study of Auction Revenue Rankings: The Case of Municipal Bonds. University of British Columbia. Discussion Paper 03-13. (forthcoming in the Rand Journal of Economics)

Silverman, B. W. (1986): Density Estimation for Statistics and Data Analysis. Monographs on Statistics and Applied Probability 26. Chapman and Hall, London.

Valkama, P. - Finkkilä, J. (2003): Pääkaupunkiseudun bussiliikenteen kilpailuttamisen talousvaikutukset. Tampereen yliopisto. Hallintotieteen laitos. Sarja A/32.

Wilson, R. (1977): A Bidding Model of Perfect Competition. Review of Economic Studies 44, 511-518.

YTV Transport Department (2001): Competitive tendering of bus services in the Helsinki Metropolitan Area 1994-2001.

Appendix 1. The theoretical model and structural estimation

The tests in HHS are based on comparing the distributions of estimated bidders' expected costs conditional on winning the auction, called pseudo-costs. The estimation of these costs is done following LPV and CPV. The models are translated to a procurement setting. In notation I use only probabilities because that is the easiest way to present the procurement version of CPV. The analysis of these models is driven by the affiliation assumption (see MW). Roughly, affiliation means that a high value of one bidder's cost estimate makes high values of other bidder's cost estimates more likely. Thus bidders would adjust their behavior if they learnt other bidders' signals.

LPV and CPV use a special case of the affiliated values (AV) model that was introduced by Wilson (1977) and generalized by MW. The restricted model is called the affiliated private values (APV) model. The AV model incorporates the entire range of models that HHS include in their testing framework. The estimation of pseudo-costs is however done using the APV model. It is still plausible to discuss things in the more general AV framework, as HHS do, because of the following discussion in LPV, page 175: "Laffont and Vuong (1996) noted that any AV model is observationally equivalent to some APV model. Thus there is no loss in explaining bids by restricting the set of models to APV models. In particular as noted by Laffont and Vuong (1996), any pure common value model is observationally equivalent to an APV model. In this sense the APV model does not exclude a priori the common value model ... APV model constitutes the most general framework identified from observed bids". Here I discuss the estimation in the APV setting. In the text I follow HHS and use the AV terminology.

The symmetric case

This model is a translation from LPV. Consider an auction in which n symmetric risk-neutral bidders compete for a single procurement contract. In the AV model the cost of fulfilling a contract for bidder i is $C_i = c_i(S, v_i)$ where v_i denotes the private signal and S is a common component. In the APV model $c_i(S, v_i) = v_i$. Private costs v_i are affiliated. Competitive behavior is identified with symmetric Bayesian Nash equilibrium strategies $s_i(v_i)$ which are increasing and differentiable. Bidder i chooses bid b_i to maximize expected profits conditional on own information v_i.

(A1.1) $\quad \max_{b_i}(b_i - v_i)\Pr(y_i \geq s^{-1}(b_i)|v_i),$

where $y_i = \min_{j \neq i} v_j$. It can be shown that the first order condition for equilibrium is

(A1.2) $\quad \Pr(y_i \geq s^{-1}(b_i)|v_i) + (b_i - v_i)[\frac{\delta \Pr(y_i \geq s^{-1}(b_i)|v_i)}{\delta y_i} \frac{1}{s\'(s^{-1}(b_i))}] = 0$

This condition is sufficient for estimation. Guerre at al. (2000) provide the identification result used in LPV. It based on the strict monotonicity of $s()$.

(A1.3) $\quad \Pr(y_i \geq s^{-1}(b_i)|v_i) = \Pr(B_i \geq B|b_i = b)$,

where B_i denotes the lowest bid among bidder i's opponents, formally $B_i = \min_{j \neq i} b_j$. It follows that

(A1.4) $\quad \frac{\delta \Pr(B_i \geq B|b_i=b)}{\delta B} = \frac{\delta \Pr(y_i \geq s^{-1}(b_i)|v_i)}{\delta y_i} \frac{1}{s\'(s^{-1}(b_i))}$

Using these two equations, (A1.2) can be written as

(A1.5) $\quad v_i = b_i + \frac{\Pr(B_i \geq B|b_i=b)}{\delta \Pr(B_i \geq B|b_i=b)/\delta B}$

As pointed out in CPV, noting that conditioning on b disappears from (A1.5), it can be interpreted as

(A1.6) $\quad v_i = b_i - \frac{\Pr(B_i \geq b, b_i=b)}{\Pr(B_i=b, b_i=b)}$

This can be estimated using nonparametric techniques. Knowing point estimates of v_i is sufficient for the HHS testing approach.

Let h_G and h_g denote bandwidths and $K()$ a kernel function. b_{it} represents the bid made by bidder i in auction t, and B_{it} represents the lowest bid among i's opponents in auction t. In the symmetric case the pseudo-cost $\hat{v}_{it}$ are estimated with the equation A1.7. The sum from t to T goes trough all the auctions with a given n, $n \geq 2$. The estimation is conducted separately for each n. In the symmetric case I use 49 auctions. The vector $T_n\' = [6(n = 2), 19(n = 3), 11(n = 4), 11(n = 5), 2(n = 6)]$ describes the total number of auctions for each n.

(A1.7) $\quad \hat{v}_{it} \equiv \hat{\xi}(b_{it}) = b_{it} - \frac{\hat{G}(b,b)}{\hat{g}(b,b)}$, where
$\hat{G}(b,b) = \frac{1}{T \times h_G \times n} \sum_{t=1}^{T} \sum_{i=1}^{n} K(\frac{b-b_{it}}{h_G}) 1\{B_{it} > b\}$ and
$\hat{g}(b,b) = \frac{1}{T \times h_g^2 \times n} \sum_{t=1}^{T} \sum_{i=1}^{n} K(\frac{b-b_{it}}{h_g}) K(\frac{b-B_{it}}{h_g})$.

The asymmetric case

This follows CPV translated to the procurement setting. I assume two different groups: group G0 consists of the bidders that have garages near the routes and group G1 includes the bidders that have garages far from the contracted route. The bidders are assumed to be symmetric within each group. G0 consists of n_0 bidders and G1 of n_1 bidders, with $n_1 + n_2 = n \geq 2$. If either n_1 or n_0 is equal to zero, the estimation procedure reduces to the one presented above for the symmetric case. The estimation equations also simplify somewhat if bidder i is the only bidder in either of the groups. The analysis must be performed separately for each given pair (n_1, n_0) because the bidding strategy of any bidder depends on both the number and the types of his opponents. Let v_{1i} denote the costs of the bidders belonging to G1 and v_{0i} the costs of the bidders in G0. Bidders draw their costs from an n-dimensional cumulative distribution $F()$. Marginal distributions may vary across subgroups. I present the model, identification and estimation strategy here only for the group G1 as bidding strategy is analogous for group G0. One should refer to CPV for more details.

Let $y^*_{1i} = \min_{j \neq i, j \in G1} v_{1j}$ and $y_{0i} = \min_{j \in G0} v_{0j}$. Then the problem for any bidder i of type 1 can be written as

$$\text{(A1.8)} \qquad \max_{b_{1i}} (b_{1i} - v_{1i}) \Pr(y^*_{1i} \geq s_1^{-1}(b_{1i}) \text{ and } y_{0i} \geq s_0^{-1}(b_{1i}) | v_{1i}).$$

Differentiating with respect to b_{1i}, the equilibrium strategy for any G1 bidder i satisfies the first-order differential equation

$$\text{(A1.9)} \qquad \Pr(y^*_{1i} \geq s_1^{-1}(b_{1i}) \text{ and } y_{0i} \geq s_0^{-1}(b_{1i}) | v_{1i}) + (b_{1i} - v_{1i}) [\frac{\delta \Pr(y^*_{1i} \geq s_1^{-1}(b_{1i}) \text{ and } y_{0i} \geq s_0^{-1}(b_{1i}) | v_{1i})}{\delta y^*_{1i}} \frac{1}{s_1'(s_1^{-1}(b_{1i}))} + \frac{\delta \Pr(y^*_{1i} \geq s_1^{-1}(b_{1i}) \text{ and } y_{0i} \geq s_0^{-1}(b_{1i}) | v_{1i})}{\delta y_{0i}} \frac{1}{s_0'(s_0^{-1}(b_{1i}))}] = 0$$

for all v_{1i} in their support, where $b_{1i} = s_1(v_{1i})$.

The identification of the model for G1 rests on the following equality:

$$\Pr(y^*_{1i} \geq s_1^{-1}(b_{1i}) \text{ and } y_{0i} \geq s_0^{-1}(b_{1i}) | v_{1i}) = \Pr(B^*_{1i} \geq B \text{ and } B_{0i} \geq B | b_{1i} = b),$$

where B^*_{1i} and B_{0i} denote the lowest bids of bidder i's opponents, formally $B^*_{1i} = \min_{j\neq i, j\in G1} b_{1j}$ and $B_{0i} = \min_{j\in G0} b_{0j}$. Again this equality is due to the strict monotonicity of $s()$. It follows that

$$\frac{\delta \Pr(B^*_{1i}\geq B \text{ and } B_{0i}\geq B|b_{1i}=b)}{\delta B} =$$
$$[\frac{\delta \Pr(y^*_{1i}\geq s_1^{-1}(b_{1i}) \text{ and } y_{0i}\geq s_0^{-1}(b_{1i})|v_{1i})}{\delta y^*_{1i}} \frac{1}{s_1\acute{}(s_1^{-1}(b_{1i}))}$$
$$+\frac{\delta \Pr(y^*_{1i}\geq s_1^{-1}(b_{1i}) \text{ and } y_{0i}\geq s_0^{-1}(b_{1i})|v_{1i})}{\delta y_{0i}} \frac{1}{s_0\acute{}(s_0^{-1}(b_{1i}))}]$$

Using these two equations (A1.9) can be written as

(A1.10) $$v_{1i} = b_{1i} + \frac{\Pr(B^*_{1i}\geq B \text{ and } B_{0i}\geq B|b_{1i}=b)}{\delta \Pr(B^*_{1i}\geq B \text{ and } B_{0i}\geq B|b_{1i}=b)/\delta B}$$

Again this can be estimated using nonparametric techniques. Below I assume that bidder i belongs to group G1 and that $n_1 \geq 2$ and $n_0 \geq 1$.

As pointed out in CPV, first note that conditioning on b_1 disappears from (A1.10) and that it can be interpreted as

(A1.11) $$v_{1i} = b_{1i} - \frac{\Pr(B^*_{1i}\geq b \text{ and } B_{0i}\geq b, b_{1i}=b)}{\Pr(B^*_{1i}=b \text{ and } B_{0i}\geq b, b_{1i}=b)+\Pr(B^*_{1i}\geq b \text{ and } B_{0i}=b, b_{1i}=b)}$$

The numerator can estimated non-parametrically by $\hat{G}_1(b_1, b_1, b_1)$. The denominator is estimated by the sum of $\hat{D}_{11}(b_1, b_1, b_1)$ and $\hat{D}_{12}(b_1, b_1, b_1)$. The sum from t to T goes trough the given pair (n_1, n_0) of the two bidder types. b_{1it} represents the bid made by bidder i of type 1 in auction t, B^*_{1it} and B_{0it} denote the lowest bids of bidder i's opponents in auction t, formally $B^*_{1it} = \min_{j\neq i, j\in G1} b_{1jt}$ and $B_{0it} = \min_{j\in G0} b_{0jt}$.

(A1.12) $$\hat{v}_{1it} = b_{1it} - \frac{\hat{G}_1(b_1,b_1,b_1)}{\hat{D}_{11}(b_1,b_1,b_1)+\hat{D}_{12}(b_1,b_1,b_1))}, \text{ where}$$

$\hat{G}_1(b_1, b_1, b_1) = \frac{1}{T\times h_{G1}\times n_1}\sum_{t=1}^T\sum_{i=1}^{n_1} 1\{B^*_{1it}\geq b_1\}1\{B_{0it}\geq b_1\}K(\frac{b_1-b_{1it}}{h_{G1}})$,

$\hat{D}_{11}(b_1, b_1, b_1) = \frac{1}{T\times h_{g1}^2\times n_1}\sum_{t=1}^T\sum_{i=1}^{n_1} K(\frac{b_1-B^*_{1it}}{h_{g1}})1\{B_{0it}\geq b_1\}K(\frac{b_1-b_{1it}}{h_{g1}})$ and

$\hat{D}_{12}(b_1, b_1, b_1) = \frac{1}{T\times h_{g1}^2\times n_1}\sum_{t=1}^T\sum_{i=1}^{n_1} 1\{B_{1it}\geq b_1\}K(\frac{b_1-B_{0it}}{h_{g1}})K(\frac{b_1-b_{1it}}{h_{g1}})$.

Considering the choice of kernel and bandwidth, in both cases LPV is followed modified by assumption 5 of HHS. A triweight kernel $K(u) = \frac{35}{32}(1-u^2)^3 1\{\mid u \mid\leq 1\}$ is used. The choice of kernel does not have much effect in practise. For bandwidths Silverman's rule of thumb (Silverman 1986) is used. Bandwidths are of the form $h = h_g = h_G = c_G(nT)^{-1/(1+2n)}$, where $c_g = c_G = 2,978 \times 1,06 \times$(empirical std. deviation of bids). The factor 2,978 follows from the use of triweight kernel instead of the Gaussian kernel. For the asymmetric the case I only use 10 auctions with $(n_0 = 2, n_1 = 3)$, 6 auctions with $(n_0 = 2, n_1 = 2)$, 5 auctions with $(n_0 = 1, n_1 = 2)$ and 8 auctions with $(n_0 = 2, n_1 = 1)$.

Appendix 2. Testing

Tests based on means

HHS propose to use the sample analog of the quantile trimmed mean, $\hat{\mu}_{n,\tau}$, as a basis of the test. Let $\hat{b}_{\tau,n}$ denote the τth quantile of observed bids. Then

(A2.1) $\hat{\mu}_{n,\tau} \equiv \frac{1}{T_n \times n} \sum_{t=1}^{T} \sum_{i=1}^{n} \hat{v}_{it} 1\{\hat{b}_{\tau,n} \leq b_{it} \leq \hat{b}_{1-\tau,n}, n_t = n\}$.

(A2.2) H_0(PC with symmetry, PC or CC with asymmetry): $\mu_{\underline{n},\tau} = ... = \mu_{\bar{n},\tau}$

H_1(CC): $\mu_{\underline{n},\tau} < ... < \mu_{\bar{n},\tau}$

To test this hypotheses one needs the variances of $(\hat{\mu}_{\underline{n},\tau}, ..., \hat{\mu}_{\bar{n},\tau})'$. Denote them with $(\frac{1}{a_{\underline{n}}}, ..., \frac{1}{a_{\bar{n}}})'$ and let $\sum$ denote the diagonal covariance matrix. HHS encourage to estimate these by a block bootstrap procedure where one auction is one block. HHS define the test statistic as

(A2.3) $\bar{\chi}^2 = \sum_{n=\underline{n}}^{\bar{n}} a_n(\mu^*_{n,\tau} - \bar{\mu})^2$, where

(A2.4) $\bar{\mu} = \dfrac{\sum_{n=\underline{n}}^{\bar{n}} a_n \hat{\mu}_{n,\tau}}{\sum_{n=\underline{n}}^{\bar{n}} a_n}$ and

$\mu^*_{\underline{n}}, ..., \mu^*_{\bar{n}}$ denotes the solution to

(A2.5) $\min_{\mu_{\underline{n}},...,\mu_{\bar{n}}} \sum_{n=\underline{n}}^{\bar{n}} a_n(\hat{\mu}_{n,\tau} - \mu_n)^2$ s.t. $\mu_{\underline{n}} \leq ... \leq \mu_{\bar{n}}$.

HHS point out that the solution to (A2.5) can be found using the "pool adjacent violators" algorithm (Ayer at al. (1955)), using the weights a_n. The p-values are then calculated using equation (A2.6) from HHS corollary 2. It states that under the null, $\bar{\chi}^2$is asymptotically distributed as mixture of Chi-squared random variables. In practise HHS suggest obtaining the weights $w(k; \Sigma)$ by simulation from the MVN$(0, \hat{\Sigma})$ distribution where estimated weights of the chi-squared-bar are defined by the distribution of the number of activated constraints in (A2.5).

Under the null PV hypothesis

(A2.6) $$Pr(\bar{\chi}^2 \geq c) = \sum_{k=2}^{\bar{n}-\underline{n}+1} \Pr(\chi^2_{k-1} \geq c)w(k;\Sigma) \;\forall\; c > 0,$$

where χ^2_j denotes a standard Chi-square distribution with j degrees of freedom, and each mixing weight $w(k;\Sigma)$ is the probability that the solution to (A2.5) has exactly k distinct values when the vector $\{\hat{\mu}_{\underline{n},t}, ..., \hat{\mu}_{\bar{n},t}\}$ has a multivariate $N(0,\Sigma)$ distribution.

A Kolmogorov-Smirnov type test

The second testing approach in HHS uses a smoothed sum of supremum distances between successive empirical distributions of pseudo-costs. Hence it is called a sup-norm test.

(A2.7)

$$\bar{\delta}_T = \sum_{n=\underline{n}}^{\bar{n}-1} \sup_{v\epsilon[\underline{v},\bar{v}]} \left\{ \begin{array}{c} \frac{1}{nT_n}\sum_{t=1}^{T}\sum_{i=1}^{n} 1\{n_t = n\}\Lambda(\hat{v}_{it} - v) \\ -\frac{1}{(n+1)T_{n+1}}\sum_{t=1}^{T}\sum_{i=1}^{n+1} 1\{n_t = n+1\}\Lambda(\hat{v}_{it} - v) \end{array} \right\}, \text{ where}$$

$\Lambda(\hat{v}_{it} - v) = \frac{\exp((v-\hat{v}_{it})/h')}{1+\exp((v-\hat{v}_{it})/h')}$ and $h' = 0.01$.

HHS suggest normalizing this distribution. I use Silverman's rule of thumb bandwidths h, defined above. HHS show that this test statistic can be approximated with subsampling. They also suggest the use of a recentering approach by Chernozhukov and Fernandez-Val (2005) where in each subsample, the test statistic is recentered by the original full-sample test statistic. Then the p-value is computed as

(A2.8) $\frac{1}{S}\sum_{s=1}^{S} 1\left\{L^s > \sqrt{T}\sum_{n=\underline{n}}^{\bar{n}-1}\sqrt{h_T(n)}\sup_x\left[\hat{F}_n(x) - \hat{F}_{n+1}(x)\right]\right\}$, where L^s is the following modified test statistic

$$L^s = \sqrt{R}\left\{ \begin{array}{c} \sum_{n=\underline{n}}^{\bar{n}-1}\sqrt{h_R(n)}\sup_x\left[\hat{F}^s_n(x) - \hat{F}^s_{n+1}(x)\right] - \\ \sum_{n=\underline{n}}^{\bar{n}-1}\sqrt{h_R(n)}\sup_x\left[\hat{F}_n(x) - \hat{F}_{n+1}(x)\right] \end{array} \right\}.$$

$s = (1, ..., S)$ refers to a particular subsample. $\hat{F}^s_n(x)$,$\hat{F}^s_{n+1}(x)$,$\hat{F}_n(x)$ and $\hat{F}_{n+1}(x)$ note the smoothed functions presented in equation (A2.7). T is the full sample size and R is the subsample size.

Appendix 3. Testing and robustness checks for the standard testing specification.

Table A3.1 shows the relationship between the p-value and the subsample size. The same subsample size is used for the distributions of the four and five bidder auctions in testing because there is the same number (11) of both these auctions. Their subsample size is presented in rows. For the 19 three bidder auctions a different subsample size is used. This subsample size is presented in columns. The numbers in the table represent the p-value from the sup-norm test for given subsample sizes with 100 draws. In practise this robustness check was first conducted with a larger region and to save computer time with a smaller number of subsamples taken (100 subsamples). All the robustness checks conducted for this study took about three weeks on a 1300 MHz computer. The relevant region was then checked again with a larger number of subsamples (500 subsamples). The results are robust for a large range of subsample sizes. Only the two smallest and two largest possible sizes from both columns and rows show nonrobust p-values. I choose subsample sizes of 9 for the three bidder auctions and 6 for the four and five bidder auctions.

Table A3.1. Robustness check for the subsample size. P-values for the sup-norm test with subsamples for three bidder auctions in columns and for four and five bidder auctions in rows. 100 subsamples taken.

Subsample size for 4-5 bidder auctions \ Subsample size for 3 bidder auctions	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	mean
2	0,03	0,00	0,02	0,01	0,04	0,03	0,00	0,02	0,07	0,09	0,07	0,06	0,06	0,11	0,12	0,14	0,16	0,06
3	0,06	0,02	0,01	0,02	0,00	0,02	0,02	0,01	0,03	0,01	0,02	0,03	0,03	0,07	0,00	0,04	0,05	0,03
4	0,06	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,02	0,01	0,03	0,01	0,01	0,00	0,01	0,00	0,02	0,01
5	0,03	0,05	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01
6	0,01	0,06	0,03	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01
7	0,05	0,01	0,00	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
8	0,06	0,05	0,04	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01
9	0,11	0,07	0,02	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01
10	0,08	0,02	0,02	0,03	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01
mean	0,05	0,03	0,02	0,01	0,01	0,01	0,00	0,00	0,01	0,01	0,01	0,01	0,01	0,02	0,01	0,02	0,03	

Table A3.2 shows the relationship between the number of subsamples taken and the p-value. The smaller the amount of subsamples taken, the easier it is to get a wrong result by chance. In this case the results are robust. To be on the safe side I choose to take 500 draws. As the p-values are now very small, enough subsamples should be taken to get accuracy in interpreting the results, especially the result zero. With 500 draws, the p-value zero really means that the p-value is smaller than $1/500 = 0.002$.

Table A3.2. Robustness check for the number of subsamples taken. Subsamples size for three bidder auctions is 9 and for the rest 6.

# of samples	**p-value**
10	0
50	0
150	0.006667
200	0
300	0
400	0.005
500	0
1000	0
2000	0.0025
5000	0.0014

In table A3.3 a summary of the results of conducting the test 50 times are presented. This was done using the subsample size nine for the three bidder auctions and subsample size six for the four and five bidder auctions. The number of subsamples taken was 500. Typically this test is just conducted once. However in some cases there can be qualitative changes in the results from one test draw to another. Lacking the proper way to choose the correct p-value a cautious researcher should choose the maximum p-value. Another possibility is to choose the mean. That is the same as conducting the test once with a larger amount of subsamples taken. In this case with 50 times 500, that is 25000 draws.

Table A3.3. Robustness check for repeating the test. Descriptive statistics of the p-value for 50 repetitions of the test with a given number of subsamples taken (500 draws) and given subsamples sizes (9 for the three bidder auctions and 6 for the rest).

	min	**max**	**mean**	**sd**
p-value	0.0000	0.0060	0.0011	0.0015

Appendix 4. Testing and robustness checks for the standard testing specification with pooling.

Table A4.1 shows the relationship between the subsample size and the p-value. Differing from the standard testing specification without pooling, the same subsample sizes are used for both distributions because the number of auctions in the two pools are almost equal ($T_{n\leq 3} = 25$ and $T_{n\geq 4} = 24$). As subsample size increases the p-value usually decreases. The inference is robust for most subsample sizes. The robustness was again first checked for the entire region with 100 subsamples taken and then again for a smaller region with 500 subsamples taken.

Table A4.1. Robustness check for p-values for the choice of subsample size with 100 and 500 subsamples taken.

Subsample size	100 draws	500 draws
2	0.00	
3	0.03	
4	0.02	
5	0.01	0.016
6	0.01	0.022
7	0.02	0.008
8	0.01	0.006
9	0.02	0.012
10	0.01	0.002
11	0.00	0.004
12	0.00	0.000
13	0.00	0.000
14	0.01	0.000
15	0.00	
16	0.00	
17	0.00	
18	0.00	
19	0.00	
20	0.00	
21	0.00	
22	0.00	
23	0.00	

Table A4.2. shows the relationship between the number of subsamples taken and the p-value with the subsample size being 9. The results are robust to the number of subsamples taken.

Table A4.2. Robustness check for the number of subsamples taken. Subsample size is 9.

# of samples	p-value
10	0
50	0
150	0.00666667
200	0
300	0.00666667
400	0.0025
500	0.002
1000	0.004
2000	0.006
5000	0.0064

In table A4.3 a summary of the results of conducting the test for 50 times are presented. There are no qualitative changes in the results from one draw to another. Therefore the test is robust to repetitions. I choose to report the largest p-value.

Table A4.3. Robustness check for repeating the test. Descriptive statistics of the p-value for 50 repetitions of the test for subsample size 9 with number of subsamples taken 500.

	min	max	mean	sd
p-value	0	0.0140	0.0063	0.0034

Appendix 5. Testing and robustness checks for the asymmetric testing specification with pooling.

Table A5.1 checks the robustness of p-value for the subsample sizes when taking 500 draws for the group with small distances. Subsample sizes for pool 3 (s=2 and l=3) are presented in rows and subsample sizes for pool 4 in columns. The results are robust for changes in subsample sizes. I choose to use 5 as the subsample size for pool 3 and 9 as the subsample size for pool 4.

Table A5.1. Robustness check for the subsample sizes with 500 subsamples taken for the group with small distances. Subsample size for auctions with less bidders in columns and subsample size for more bidders in rows.

Subsample size for pool 3	Subsample size for pool 4							
	5	**6**	**7**	**8**	**9**	**10**	**11**	**12**
3	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0
5	0.002	0	0	0	0	0	0	0
6	0.002	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0

Table A5.2. shows the relationship between the number of subsamples taken and the p-value with the subsample size for pool 3 being 5 and the subsample size for pool 4 being 9 for the group with small distances. Again the results seem to be robust for the entire range of subsamples taken.

Table A5.2. Robustness check for the number of subsamples taken for the group with small distances. Subsample size for pool 3 = 5 and subsample size for pool 4 = 9.

# of samples	p-value
10	0
50	0
150	0
200	0
300	0
400	0
500	0
1000	0
2000	0
5000	0

In table A5.3 a summary of the results of conducting the test for 50 times are presented. There are no qualitative changes in the results from one draw to another. Therefore the test is robust to repetitions. As before the maximum is reported.

Table A5.3. Robustness check for repeating the test for the group with small distances. Descriptive statistics of the p-value for 50 repetitions of the test for subsample size 3 = 5 and subsample size 4 =9 with number of subsamples taken 500.

	min	max	mean	sd
p-value	0.0000	0.0020	0.0000	0.0003

The first check for the large distance group is presented in table A5.4. Subsample sizes for pool 1 (more bidders) are presented in rows and subsample sizes for pool 2 (less bidders) in columns. The results seem to be robust for changes in subsample sizes and no qualitative changes exist. As subsample sizes increase the p-value decreases. I choose to use 5 as the subsample size for pool 1 and 9 as the subsample size for pool 2.

A5.4 Robustness check for the subsample sizes with 500 subsamples taken for the group with large distances. Subsample size for less bidders in columns and subsample size for more bidders in rows.

Subsample size for pool 1	Subsample size for pool 2: 5	6	7	8	9	10	11	12	mean
3	0.56	0.55	0.54	0.52	0.50	0.57	0.55	0.56	0.54
4	0.51	0.43	0.52	0.50	0.48	0.49	0.52	0.54	0.50
5	0.53	0.49	0.51	0.48	0.47	0.44	0.48	0.47	0.48
6	0.51	0.49	0.48	0.44	0.45	0.42	0.43	0.40	0.45
7	0.51	0.48	0.47	0.41	0.43	0.41	0.39	0.39	0.44
mean	0.53	0.49	0.50	0.47	0.47	0.47	0.47	0.47	

The results shown in table A5.5 seem to be fairly robust for the entire range of the number subsamples taken.

Table A5.5. Robustness check for the number of subsamples taken for the group with large distances. Subsample size 1 = 5 and subsample size 2 =9.

# of samples	p-value
10	0.40
50	0.46
150	0.43
200	0.43
300	0.47
400	0.49
500	0.46
1000	0.45
2000	0.47
5000	0.46

Table A5.6 shows that in this case there are no qualitative changes in the results from one test draw to another. Therefore the test is robust to repetitions. As before the maximum is reported.

Table A5.6. Robustness check for repeating the test for the group with large distances. Descriptive statistics of the p-value for 50 repetitions of the test for subsample size 1 = 5 and subsample size 2 =9 with number of subsamples taken 500.

	min	max	mean	sd
p-value	0.44	0.53	0.47	0.02