Combination Bidding in Multi-Unit Auctions*

Estelle Cantillon† and Martin Pesendorfer‡

First Version: July 2001
This Draft: November 2006

Abstract

This paper considers the problem of identification and estimation in the first-price multi-unit auction. It is motivated by the auctions of bus routes held in London where bidders submit bids on combinations of routes as well as on individual routes. We show that submitting a combination bid lower than the sum of the bids on the constituent routes does not require cost synergies and can instead serve as a tool to leverage market power across the different routes. As a result, the welfare consequences of allowing combination bidding in the first price auction are ambiguous, and depend on the importance of the cost synergies. We provide conditions for non-parametric identification of the multidimensional private information in the multi-unit first price auction and derive partial identification results when they are not satisfied. We propose an estimation method consisting of two stages: In the first stage, the distribution of bids is estimated parametrically. In the second stage, the (set of) costs and distribution(s) of costs consistent with the observed behavior are inferred based on the first order conditions for optimally chosen bids. We apply the estimation method to data from the London bus routes market. We quantify the magnitude of cost synergies and assess possible efficiency losses arising in this market.

*We are grateful to several officials at London Transport Buses and at the London Omnibus Traction Society for their help in collecting the data, and to Ioannis Ioannou, Bryan Lincoln, Bill Simpson, Qinan Tang and Jay Cox for their excellent research assistance. We are indebted to many seminar and conference audiences and to many individuals, in particular Mireia Jofre-Bonet, Ariel Pakes, Wolfgang Pesendorfer and Bill Simpson, for their helpful suggestions. Costas Meghir and four anonymous referees made comments that substantially improved this draft. We thank the NSF under grant SES 0214222 and the Division of Research at Harvard Business School for financial support.

†FNRS, ECARES (Université Libre de Bruxelles) and CEPR. Email: Estelle.Cantillon@ulb.ac.be

‡Department of Economics, London School Economics, CEPR and NBER. Email: m.pesendorfer@lse.ac.uk.

1 Introduction

This paper considers the problem of identification and estimation in the first-price multi-unit auction. It is motivated by the auctions held by the London Transportation authority to award contracts to service bus routes. Two special features of these auctions are that several bus routes are auctioned at the same time, and that bidders may submit combination bids in addition to stand-alone bids. In other words, the London bus routes market is an example of a combinatorial auction.

Combinatorial auctions allow bidders to transmit rich information regarding the value they attach to the objects for sale. When bidders value bundles of objects differently than the sum of the constituent parts, allowing bids for a combination of routes is a necessary condition for efficiency and optimality. This was well understood by the London Transportation authority. Indeed, two of the motivations for allowing combination bids in the London bus routes market were: (i) that they would allow bidders to pass on some of the cost savings resulting from cost synergies between routes through lower bids, and (ii) that they would enhance the efficiency of the allocation of routes across bidders.

However, allowing combination bids in the first price auction may also hurt efficiency and costs. Section 2 introduces a model of a private value multi-unit first price procurement auction that allows for cost synergies between routes. Two distinct motivations for combination bidding are illustrated. First, combination bidding gives rise to a strategic effect because bidders' stand-alone bids compete with their combination bids. As a result, bidders may find it profitable to inflate their stand-alone bids relative to their combination bids in order to favor the latter, even in the absence of any cost synergy. Second, when cost synergies are important, the fact that combination bids do allow bidders to align their bids better on their costs can help improve efficiency and lower costs. As a result, the welfare consequences of combination bidding depend on which motivation dominates.

How to disentangle the alternative motivations for combination bidding is an open empirical question. This paper contributes to this question. In doing so, we follow the footsteps of the seminal papers by Elyakime, Laffont, Loisel and Vuong (1994) and Guerre, Perrigne and Vuong (2000) who introduced an indirect approach to the empirical analysis of auctions. The earlier "direct" approaches include those proposed by Donald and Paarsch (1993, 1996), Laffont and Vuong (1993) and Laffont, Ossard and Vuong (1995). These approaches are all intrinsically parametric. The primitives, including the distribution of costs, are given specific parametric forms. The estimation proceeds by choosing the value of the parameters for which the equilibrium distribution of bids is as close as possible to the empirical distribution of bids. These approaches require solving for the equilibrium, which can be difficult and time-consuming.

Elyakime, Laffont, Loisel and Vuong (1994) and Guerre, Perrigne and Vuong (2000) show that bidders' first order conditions for optimally chosen bids provide an expression for their private information as a function of their bids and the distribution and density of equilibrium bids. Since the bids are observable, the theoretical expressions involving the distribution and density of bids can be replaced with the empirical counterparts suggesting a two-step estimation method. In the first step, the distribution and density of bids are estimated. In the second step, the private information and the bidders' distribution of private information is estimated as a residual in the bidders' first order condition. The method does not require the computation of equilibrium strategies.

The indirect approach has boosted research on non-parametric identification of auction models (see Athey and Haile, 2004 for a survey of recent results) and, because the estimation does not rely on the computation of the equilibrium bid functions, it has led to a rapid development of auction applications beyond the single object symmetric independent private value model (e.g. Li, Perrigne and Vuong, 2002, Campo, Perrigne and Vuong, 2003, Jofre-Bonet and Pesendorfer, 2003 and Athey, Levin and Seira, 2004).

Section 3 studies our identification problem. We assume that bidders assess their winning probabilities correctly and choose bids to maximize their profits. We also assume that the observed data capture all the relevant characteristics of the environment. The model is said to be (non parametrically) identified under these assumptions, if the primitives, i.e. the costs and the distribution of costs, can be uniquely inferred from the observed data.

As in Guerre, Perrigne and Vuong (2000), our identification arguments are based on the best response conditions given by the first order conditions for optimal bids. Yet, there are notable differences between our identification problem and theirs.[1] First, Guerre, Perrigne and Vuong build on the known characterization of equilibrium in the single object independent private value symmetric first price auction to argue that the first order condition for optimal bids identifies costs, for all bids above the reserve price. In our case, no such characterization result is available. We prove a new partial characterization result for the combinatorial first price auction that justifies when the first order conditions can be used for identification (Lemma 1). In some cases, optimal bids satisfy a set of inequalities rather than a set of equalities and thus first order conditions do not completely identify costs.

[1] Extensions of Guerre, Perrigne and Vuong's results and the relationship between our results and those are discussed below under "related literature."

Second, optimal bids are characterized by a single first order condition in the single object auction, whereas they are characterized by a *system* of equations in our case. The consequence is that identification requires an extra step to prove that this system of equations admits a unique solution (Lemma 2). Third, in the single object independent private value model, the main source of under-identification is the presence of a reserve price. In our case, there are additional sources of underidentification. We address each of these systematically and characterize what can be identified in each case.

Our results are as follows. The multi-unit first price auction model is identified when bidders actively bid, that is, submit bids that have a strictly positive probability of winning, on all contracts and combinations thereof. In practice however, several factors may reduce the dimensionality of the observed information. First, not submitting a bid that has a positive probability of winning on all items can be part of equilibrium behavior. Second, a constraint on bids, such as a reserve price, or the condition that combination bids must be lower than the sum of the constituent bids, can further reduce the dimensionality of the observed information. In these cases, partial identification emerges and a (non singleton) set of costs and cost distributions is identified.

Section 4 turns to our application and describes the London bus routes market. This market is particularly well-suited for this kind of analysis. First, there is a common perception that synergies between routes are prevalent. Second, combination bids are permitted and play an important role in this market with about 30% of all routes won by combination bids. Thus, our method allows us to quantify the extent of cost synergies in this market, and therefore assess the role of combination bids.

Section 5 proposes our estimation method. The estimation proceeds in two stages. In the first stage, the multi-variate joint distribution of bids for all units is estimated. Due to the multi-dimensionality of the bid vectors, and to incorporate covariates, we consider a parametric bid density. In the second stage, the costs or the set of costs that rationalize the bids are inferred by using the first order conditions for optimal bids.

Section 6 reports our estimates. We illustrate the bidders' objective function by using the estimated winning probability. The estimates suggest that bidders may find it profitable to submit a combination bid less than the sum of stand-alone bids even in the absence of cost synergies. We report our bound estimates on possible cost synergies. We find little evidence of cost synergies. Our estimates suggest that a route combination costs more than the sum of stand-alone costs on average. A possible explanation is that the technology of operating buses in London depends on the garage capacity and the number of garages a bidder has available. As the number of buses in the fleet approaches the available garage capacity it becomes increasingly costly to take on additional buses or routes. At the end of this section, we use our cost estimates to assess bounds on possible efficiency losses.

Related literature. There is a growing literature on identification and estimation in auctions. Donald and Paarsch (1993), Laffont, Ossard and Vuong (1995), Guerre, Perrigne and Vuong (2000) and others propose identification results and estimation techniques to infer bidders' private information (Athey and Haile, 2004, provide a survey of identification results in auctions). The literature focuses to a large extent on the single-unit auction model and little is known about auctions in which multiple units are sold. Exceptions include the sequential auctions analyzed by Jofre-Bonet and Pesendorfer (2003) and Donald, Paarsch and Robert (2005) and the simultaneous auctions of homogenous goods analyzed, among others, by Hortacsu (2002), Wolak (2003), Fevrier, Preget and Visser (2004), Hortacsu and Puller (2005) and Kastl (2005). Jofre-Bonet and Pesendorfer propose an estimation techniques to measure linkages between items sold at sequential auctions. They find evidence of substitutes at sequential highway procurement auctions. Hortacsu (2002) and Fevrier, Preget and Visser (2004) study share auctions for treasury bills, and Wolak (2003) and Hortacsu and Puller (2005) study electricity auctions. Assuming that the equilibrium bid functions are strictly decreasing demand functions and generate a residual supply for bidders that has no kinks, Hortacsu (2002) shows that the private value multi-unit discriminatory auction model is non parametrically identified. Wolak (2003) discusses in more details how constraints on bids, such as the requirement that bids belong to a price grid, affect identification. Finally, McAdams (2005) has recently argued that the homogeneous good multi-unit auction model is only partially identified in general, and that previous approaches relied on more or less implicit identification assumptions on optimal bids. The identification problem in share auctions is close to ours in the sense that the model is also multi-dimensional and identification proceeds via the first order condition of optimal bids. Yet, a key difference between our setting with heterogeneous goods and the share auction model with homogeneous and divisible goods is that demand is identified by a *vector* of costs $(c_1, ..., c_S) \in \mathbb{R}^{2^m-1}$ in our setting whereas it is identified by a marginal valuation *function* in share auctions. This leads to different mathematical structures.

The first order condition of equilibrium prices as an estimation condition features dominantly in the empirical literature on differentiated products. Goldberg (1995) incorporates quota constraints into the estimating equations, which result in Kuhn-Tucker conditions as in our framework. A distinguishing feature of our auction problem is that bidders have incomplete information about opponents' costs and bids. In Goldberg, firms have complete information about opponents' costs and prices leading to distinct best response conditions.

There has also been a number of recent theoretical analyses of auctions of heterogenous objects. Among these, Armstrong (2000) and Avery and Hendershott (2000) derive properties of the optimal multi-unit auction when types are multidimensional and objects may be substitutes or complements. A central question that these authors address is to what extent the auctioneer may benefit from bundling the objects (A seminal contribution to this question is Palfrey, 1983). Krishna and Rosenthal (1995) and Branco (1997) study the second price multi-unit auction with synergies. Milgrom (2000) highlights some perverse effects of combinatorial bidding in ascending auctions. Our analysis contributes to this literature by highlighting the motivations and consequences of combination bidding in the combinatorial first price auction. The strategic motivation we uncover is analogous to the bundling motivation in the (decision-theoretic) multi-dimensional screening literature (McAfee, McMillan and Whinston, 1989, Armstrong, 1996 and Armstrong and Rochet, 1999) but it had never been pointed out in the auction context.

Finally, the importance of synergies in multi-unit auctions has been emphasized by the recent experience in FCC spectrum auctions. Ausubel, Cramton, McAfee and McMillan (1997) and Moreton and Spiller (1998) use a regression analysis to measure synergy effects in these auctions. Other recent applications of combinatorial auctions include the auctions for school meals in Chile (Epstein, Henriquez, Catalan, Weintraub and Martinez, 2002) and corporate procurement applications at Sears and Home Depot. (Cramton, Shoham and Steinberg, 2006, provide a survey of recent issues and applications in combinatorial auctions.)

2 Bidding environment

This section introduces the model and highlights its key properties.

2.1 Model

Our model integrates the salient features of the London bus routes market. A procurement agency (the "buyer") simultaneously invites bids on m routes from N risk neutral bidders. A contract covers a single route or a combination of routes. Each bidder i privately observes a cost draw, $c_s^i \in \mathbb{R}$, for each contract. We let S denote the set of routes. To simplify notations, we adopt the convention that $s \subseteq S$ means that s is a non empty subset of S.

Let $\mathbf{c}^i = (c_s^i)_{s \subseteq S}$. (Vectors and matrices are in bold, scalars are in standard fonts). We say there is no cost synergy if $c_s^i + c_t^i = c_{s\cup t}^i$ for contracts s and t with $s \cap t = \emptyset$, where $c_{s\cup t}^i$ denotes bidder i's cost for the combination of contracts s and t. Cost synergies are positive if costs are strictly subadditive, $c_{s\cup t}^i < c_s^i + c_t^i$, and negative if $c_{s\cup t}^i > c_s^i + c_t^i$.

Information. Bidders' contract costs are independently distributed according to the distribution $F_i(.|\mathbf{X}) : \mathcal{C}_i \subseteq \mathbb{R}^{(2^m-1)} \to [0,1]$, $i = 1, ...N$, where $\mathbf{X} = (\mathbf{x}, \mathbf{w})$ denotes a vector of observable contract characteristics $\mathbf{x}$ and bidder characteristics $\mathbf{w}$. We assume that F_i is common knowledge and that F_i and $\mathcal{C}_i$ satisfy the following condition:

Assumption 1: For all i, $\mathcal{C}_i$ has a non empty interior and F_i has a strictly positive and continuous density on $\mathcal{C}_i$.

Let $\mathcal{F}$ be the set of distribution functions that satisfy Assumption 1. The assumption that $\mathcal{C}_i$ has a non empty interior ensures that the distribution is not degenerate and that bidders' private information is indeed of dimension $2^m - 1$. The assumption on the density is flexible. It allows independence and correlation or affiliation in bidders' costs across contracts.

Auction rule. We consider the auction rule used in the London bus routes market. Bidders may submit bids on all contracts. Let b_s^i denote bidder i's bid on contract s, and let $\mathbf{b}^i = (b_1^i, ..., b_s^i, ..., b_S^i) \in \mathbb{R}^{2^m-1}$. Bidders pay the value of their winning bids and the buyer selects the winner(s) based on the allocation that minimizes her total payment (first price auction). In the London bus route market, the buyer imposes the additional requirement that bids be subadditive, that is, $b_{s\cup t}^i \leq b_s^i + b_t^i$ for all s, t such that $s \cap t = \emptyset$. There may also be a reserve price. Let R_s denote the reserve price for contract s. Bids on or above the reserve price are rejected. Ties are resolved by randomizing over all cost-minimizing allocations.

Payoffs. Fix bidder i, and for each contract $s \subseteq S$, define $\underline{b}_s^{-i}$ as the value of the cheapest allocation of the routes in s among bidder i's opponents. Formally, let $\mathcal{P}^{N\backslash i}(s)$ define the set of partitions of s into $N-1$ (possibly empty) subsets. Let $t = (t_1, ..., t_{i-1}, t_{i+1}, ...t_N)$ denote an element of $\mathcal{P}^{N\backslash i}(s)$. Then,

$$\underline{b}_s^{-i} = \min_{t \in \mathcal{P}^{N\backslash i}(s)} \sum_{j \neq i} b_{t_j}^j \tag{1}$$

where, by convention, $b_\emptyset^j = 0$. Note that by construction $\underline{b}_s^{-i}$ is subadditive.

Given the auction rules and bidders' risk neutrality, bidder i's payoff is given by (ignoring ties):

$$\begin{cases} b_s^i - c_s^i & \text{if } b_s^i + \underline{b}_{S\backslash s}^{-i} < \min_{t \subseteq S,\, t \neq s}\{b_t^i + \underline{b}_{S\backslash t}^{-i}, \underline{b}_S^{-i}\} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

2.2 Properties

Not much is known about the general properties of the equilibrium in this auction, beyond the fact that an equilibrium exists. When the support of costs is bounded, the game satisfies the conditions of Theorem 1 of Jackson, Simon, Swinkels and Zame (2002) for a (mixed strategy) equilibrium with endogenous tie-breaking rule. When values are private, we can argue that whenever ties occur, bidders are indifferent among the ways in which they are resolved.[2] Hence, the equilibrium of the extended game with endogenous tie-breaking rule is also an equilibrium of the original game.

Additionally, we can get some insight into the bidding behavior we may expect at equilibrium by drawing on the analogy between the bidders' optimization problem in this auction and that of the multi-product monopolist. To fix ideas consider a two-object auction (the generalization to more objects is straightforward). Consider bidder i. With two objects, there are four possible allocations between bidder i and his opponents: bidder i wins object 1 (and one of his opponents wins object 2, if $b_1^i + \underline{b}_2^{-i}$ corresponds to the cost of the cheapest allocation), he wins object 2 (and his opponents win object 1), he wins both objects or he does not win anything.

Now consider object 1. Holding the distribution of the value of the cheapest allocation among his opponents $(\underline{b}_1^{-i}, \underline{b}_2^{-i}, \underline{b}_{12}^{-i})$ fixed, decreasing b_1^i increases bidder i's chance to win exactly object 1 by lowering the price of allocation that corresponds to $b_1^i + \underline{b}_2^{-i}$ relative to the others. But it decreases bidder i's chance of winning object 2 or the package because it could be the case that, *had bidder i not lowered b_1^i*, the value of the cheapest allocation was $b_2^i + \underline{b}_1^{-i}$ or b_{12}^i. Another way to look at this is in terms of the following trade-off. The benefit to bidder i from lowering his bid on object 1 is that he wins object 1 more often. The costs are twofold. First, it lowers his profit margin whenever he wins object 1. Second, it reduces his chance of winning object 2 alone or the package of objects 1 and 2. At equilibrium, of course, bidder i chooses bid b_1^i such that this marginal benefit and these marginal costs exactly balance one another. This trade-off is at the heart of the multi-unit first price auction: bidders' own bids compete with one another.

This is also a classic price discrimination trade-off, analogous to that present in the multi-product monopolistic pricing problem. When the multi-product monopolist sets his

[2] A formal proof of this statement is provided in Lemma 1 below. The only difference with analogous arguments made for single object private value auctions (e.g. Maskin and Riley, 2000, and Jackson et al., 2002) is that, here, changing the bid on one object not only changes the probability of winning that object but also changes the probability of winning other objects.

prices, he takes into account the fact that the price of a product affects both the demand for this product and the demand for other products.

To make this analogy more transparent, consider the multi-product monopolist problem. For simplicity, suppose the monopolist is selling two products to a buyer with unit demand and private information about his value for these products, v_s, $s \in \{1, 2, 12\}$. The buyer has additive preferences, i.e. $v_1 + v_2 = v_{12}$. The monopolist sets a price, b_s, for each product and for the combination. Then, the buyer will buy product (or bundle) s if $v_s - b_s > v_t - b_t$ for all $t \neq s$ and $t \subseteq S$ and $v_s - b_s > 0$. Rearranging, and exploiting the fact that preferences are additive, yields the following payoff for the monopolist:[3]

$$\begin{cases} b_s - c_s & \text{if } b_s + v_{S\setminus s} < \min_{t\subseteq S,\, t\neq s}\{b_t + v_{S\setminus t}, v_S\} \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

Comparing this expression with that obtained for the bidder in the combinatorial auction, (2), it is clear that the bidders' optimization problem is identical to that of the monopolist facing a buyer with additive preferences.[4]

Consequently, some of the results from the literature on monopolistic multiproduct pricing apply to describe bidders' best responses in our setting. McAfee, McMillan and Whinston (1989) derive a *sufficient* condition for bundling to be profitable for the monopolist with additive costs. They find that when demand is independent across goods – in our setting, whenever the value of the cheapest allocation among bidder i's opponents are additive and independently distributed across products – submitting a price for the bundle that is lower than the sum of the individual prices is optimal for the monopolist. Schmalensee (1984) examines the special case of gaussian demands. Armstrong and Rochet (1999) solve for the global maximization of the multi-product monopolist when valuations are binary. Their analyses confirm that bundling (the equivalent of submitting a combination bid $b_{1\cup 2} < b_1 + b_2$) is profitable unless there is strong correlation across buyers' valuations.

Two lessons emerge from the multi-product monopoly pricing literature. First, observing a combination bid lower than the sum of the bids for the constituent units in our data is no guarantee that there are underlying cost synergies. Submitting a combination bid can be profitable exactly for the same reason why the multi-product monopolist finds price discrimination profitable. Section 3 explores how the additional knowledge about the envi-

[3] Consider: $v_s - b_s > v_t - b_t$ iff $b_s + v_t + v_{S\setminus t\cup s} - v_{t\cap s} < b_t + v_s + v_{S\setminus t\cup s} - v_{t\cap s}$ iff $b_s + v_{S\setminus s} < b_t + v_{S\setminus t}$.

[4] With the caveat of course that the valuations are additive in (3) whereas they are subadditive in (2). In that sense, (3) is nested in (2).

ronment a bidder was facing when submitting his bids allows us to disentangle these two motivations (cost synergies versus strategic price discrimination). Second, correlation in the values of the cheapest allocation among bidders' opponents across the different routes is an important determinant of combination bidding (Schmalensee, 1984, McAfee, McMillan and Whinston, 1989, Armstrong and Rochet, 1999). The more correlated these values are, the more likely it is that cost synergies are the main driver of combination bidding. The less correlated these bids are, the greater the motivation for strategic price discrimination.

2.3 Welfare

In the presence of cost synergies, allowing bidders to convey information about their costs for the combinations of contracts, *in addition* to their costs for the individual contracts, is a necessary condition for efficiency (Groves, 1973, Clarke, 1971) and for optimality (Levin, 1997, Armstrong, 2000 and Avery and Hendershott, 2000). In fact, this is why the procurement authority for the London bus routes allows combination bids in their tendering process.

It is straightforward to generate examples where, indeed, allowing combination bids in the multi-unit first price auction improves efficiency and lowers procurement costs. In this subsection, we illustrate through a simple example that allowing combination bids may also have costs. In the following example, an independent first price auction for each contract is efficient and minimizes the expected procurement cost. Allowing combination bids destroys efficiency and increases expected procurement costs. In the example, the cause can be traced back to the strategic (price discrimination) effect identified above.

Example 1: There are two routes and three bidders. Bidders A and B are the "local" bidders. They care only about one route. Bidder A has cost c_1^A for route 1. Bidder B has cost c_2^B for route 2. Bidder C is a global bidder. He has an interest in both routes. (In the context of our application, one can think of bidder C as having a garage between the two routes, bidder A has his garage close to route 1 but far from route 2, etc.). Bidder C's costs are given by (c_1^C, c_2^C, c_{12}^C) where $c_{12}^C = c_1^C + c_2^C$ (no cost synergy). Costs are private information and independently distributed across bidders. Moreover, assume that, at the route level, bidders are symmetric. In other words, c_1^A and c_1^C have the same ex-ante distribution, and so do c_2^B and c_2^C.

It is instructive to first consider the scenario when combination bids are not allowed, that is, the buyer holds two independent auctions, one for route 1 and the other for route 2. The allocation of each route is independent of the outcome and the bids for the other route. Since bidders are symmetric at the individual market level (only bidders A and C bid on contract 1 and only bidders B and C bid on contract 2), the unique equilibrium is symmetric and in strictly increasing strategies. Thus, the bidder winning route 1 is the bidder who has the lowest cost for route 1 and similarly for route 2: the outcome is efficient. Moreover, if private information is one-dimensional, and conditional on the optimal reserve price and the usual regularity condition, this simple auction format minimizes procurement costs (Levin, 1997).[5]

Now suppose that bidders are allowed to submit combination bids. At equilibrium, the global bidder will submit bids such that $b_{12}^C < b_1^C + b_2^C$. Indeed, consider his optimization problem. From his perspective, the values of the cheapest allocation of the routes among his opponents, $(\underline{b}_1^{-C}, \underline{b}_2^{-C}) = (b_1^A, b_2^B)$, are independently distributed (since their costs are independently distributed) and they are additive, $\underline{b}_{12}^{-C} = \underline{b}_1^{-C} + \underline{b}_2^{-C}$. Therefore, the analysis of McAfee, McMillan and Whinston (1989) applies and bidder C will find it advantageous to submit a non trivial combination bid at equilibrium, $b_{12}^C < b_1^C + b_2^C$.[6] *Combination bidding must take place in any equilibrium.* The intuition is that the combination bid pools the two contracts together and allows the global bidder to leverage any advantage he has for one contract into the other. Indeed, suppose that bidder A has a very high cost realization for route 1. Then, the global bidder has an advantage for that route. If the global bidder only submits a combination bid, he reduces the toughness of the competition he faces for the second route because bidder B needs to submit a really low bid to compensate for bidder A's high bid and have a chance to win. This mechanism, route linkage through combination bidding, is analogous to the leverage theory in industrial organization (Whinston, 1989).

In this example, combination bidding hurts efficiency because whether bidder A wins route 1 or not, no longer depends on bidder A's and bidder C's costs only, but also on bidder B's cost (through the combination bid of the global bidder).

Moreover, if private information is one-dimensional, the expected procurement cost of the combinatorial auction is also higher than the expected procurement cost resulting from two independent first price auctions. The argument for this claim relies on a version of the revenue equivalence theorem (Myerson, 1981). When bidders are risk neutral and private information is independently distributed across bidders, the expected procurement cost of

[5] The optimal auction when private information is multidimensional is not known for this setting.

[6] This argument implicitly assumes that bidder A only submits a bid on route 1 and bidder B only submits a bid on route 2 at equilibrium. There is also a (degenerate) equilibrium where bidders only submit a bid on the package (and no bids on the individual contracts). The same conclusion, inefficiency and higher costs, applies to this equilibrium.

any incentive compatible allocation mechanism is fully determined by the allocation (see Levin, 1997, for a derivation that applies to the environment here). We argued above that the expected cost in the independent auctions case was the minimum attainable (and that this is the unique allocation that reaches that minimum). We have also just argued that the allocations differ across the two auction formats. Thus, the expected procurement cost in the combinatorial auction must be strictly higher. By continuity, we can argue that if private information is multidimensional but sufficiently correlated, the independent auctions format still yields lower procurement costs than the combinatorial auction format.[7] Moreover, the result in McAfee, McMillan and Whinston extends to positive cost synergies and to low enough negative cost synergies. Thus so does the result in example 1.

Understanding the costs and benefits of allowing combination bids is an important policy question. Example 1 suggests that the question of which format is better for the London bus routes market is ultimately an empirical one because the answer depends on the nature and extent of synergies present in the market.

3 Identification

This section describes our identification results for the combinatorial first price auction. We observe data on all bids, contract characteristics and bidder characteristics. We make the following assumption on the data generation process:

Assumption 2: We observe a cross section of auction data, $(\mathbf{b}^{it}, i = 1..., N, \mathbf{X}^t)_{t=1}^T$ from a i.i.d. random sample of T independent auctions.[8] There is no unobserved contract or bidder heterogeneity, beyond the privately observed costs. The data is generated by equilibrium play in which bidders' bids are equal or above their costs (if there are multiple equilibria, we assume that the same equilibrium is played in all auctions).[9]

[7] The revenue equivalence theorem generalizes to environments with multidimensional private information (see Krishna and Perry, 2000).

[8] By convention, non submitted bids on some contracts are assigned an arbitrary large value that guarantees they never win.

[9] Equilibrium behavior requires that bidders are best responding to their opponents' equilibrium distribution of bids. The assumption that bidders play the same equilibrium in all auctions is a common, though not trivial, assumption in empirical auction works. Bidding on or above costs follows from equilibrium behavior for bids that win with positive probability, but it is not necessary for bids that have zero probability of winning. This technical assumption is needed for the proof of Lemma 1 but is otherwise inessential for the identification argument. Any equilibrium strategy where a bidder bids below his costs can be replaced by an equilibrium strategy where these bids are at or above costs without affecting bidders' payoffs.

The model is said to be (non parametrically) identified if the primitives, i.e. the distributions of costs $F_i(.|\mathbf{X}) \in \mathcal{F}$, can be uniquely inferred from the observed data under the assumption that bidders play equilibrium strategies and that the observed data capture all the relevant characteristics of the environment. The model is said to be partially identified if F_i can be shown to belong to a proper subset of $\mathcal{F}$ for all i.

In this section we first discuss the nature of the identification problem in the multi-unit first price auction and derive the needed equilibrium characterization results.

We then derive the conditions under which the model is identified. We find that the combinatorial first price auction model is identified if, at equilibrium, bidders submit bids that have a strictly positive probability to win all contracts (Theorem 1). Unfortunately, we cannot a priori rule out irrelevant bids (i.e. bids that never win) as part of equilibrium behavior. In section 3.2, we show that irrelevant bids on some contracts do not affect the identification of the costs of the other contracts. They remain identified. By contrast, only bounds on the cost of the contract that received the irrelevant bid are identified. The potential underidentification of the costs associated to some bid realizations implies that the distribution of costs is only partially identified. A set of cost distributions rather than a unique distribution of costs rationalizes the observed bids. However, we argue that the implied bounds on the distribution can still be used for policy analysis.

Real-life combinatorial auctions may of course differ from the benchmark combinatorial auctions by introducing further constraints on the set of observable bids. In section 3.3, we examine two such constraints that arise in our application: reserve prices and the constraint that bids on combinations be lower than the sum of the bids on the constituent parts. Both constraints introduce an additional level of underidentification. Again, we derive bounds on the costs and cost distributions when these constraints bind. However, unlike before, the implied bounds on the distributions of costs do not, in general, provide bounds on policy outcomes of interest. All distributions of costs consistent with equilibrium behavior must be used to evaluate such policy outcomes.

3.1 Preliminaries

We start with the general (unconstrained) combinatorial first price auction. Bidders submit bids on all contracts, the auctioneer selects the cheapest bidder-bid allocation and the winners pay the price of their winning bids. Ties are resolved by randomizing over all cost-minimizing allocations..

Fix bidder i and let $\underline{\mathbf{b}}^{-i}$ be the $2^m - 1$ dimensional vector of the values of the cheapest allocation of each bundle among bidder i's opponents (this was defined formally in (1)). Given m routes, there are 2^m possible allocations of the contracts between bidder i and his opponents. Let $G_s(.|\mathbf{X})$: $\mathbb{R}^{2^m-1} \rightarrow [0,1]$ denote the probability that a bid vector by bidder i wins exactly contract s, conditional on some covariates $\mathbf{X}$. The functions G_s are non parametrically identified from the data when all submitted bids and all relevant contract characteristics are observed. (In the following, we simplify notation by dropping the $\mathbf{X}$ arguments in the G_s functions. They do not play any role in the identification.)

In the combinatorial first price auction, bidders solve the following $2^m - 1$ dimensional optimization problem:

$$\pi^i(\mathbf{c}^i) = \max_{\mathbf{b}^i \in \mathbb{R}^{2^m-1}} \sum_{s \subseteq S} (b_s^i - c_s^i) G_s(\mathbf{b}^i) \tag{P1}$$

Each function G_s is monotonically decreasing in b_s^i (this makes bidder-bid combination $b_s^i + \underline{b}_{S \backslash s}^{-i}$ more expensive) and monotonically increasing in b_t^i for $t \neq s$, so G_s is differentiable almost everywhere (a.e.) and with a bounded derivatives a.e. (Billingsley, 1995, thm 31.2). Therefore, the objective function in (P1) is also a.e. differentiable.

However, a.e. differentiability of the objective function in (P1) is not enough to use the first order conditions for optimal bids as part of our identification strategy. To clarify the nature of the problem, define as $\mathcal{B}_i$ the set of bidder i's equilibrium bids. It can be partitioned into three subsets:

1. $\mathcal{B}_{i1} = \{\mathbf{b}^i \in \mathcal{B}_i$: the objective function in (P1) is differentiable at $\mathbf{b}^i$ and the first order conditions of (P1) identify a unique cost vector $\mathbf{c}^i\}$
2. $\mathcal{B}_{i2} = \{\mathbf{b}^i \in \mathcal{B}_i$: the objective function in (P1) is differentiable at $\mathbf{b}^i$ but the first order conditions of (P1) do not identify a unique cost vector $\mathbf{c}^i\}$
3. $\mathcal{B}_{i3} = \{\mathbf{b}^i \in \mathcal{B}_i$: the objective function in (P1) is not differentiable at $\mathbf{b}^i\}$

Let $\eta^i(.)$ denote the probability measure of bidder i's equilibrium bids. Let $\mu(.)$ denote the Lebesgue measure. The model is identified if we can associate a density to almost every cost in $\times_i \mathcal{C}_i$. Given the definition of $\mathcal{B}_{i1}, \mathcal{B}_{i2}$ and $\mathcal{B}_{i3}$, this will be trivially the case if $\eta^i(\mathcal{B}_{i1}) = 1$. The model is partially identified otherwise.

These notations allow us to compare the nature of the identification problem here and that in Guerre, Perrigne and Vuong (2000). Guerre, Perrigne and Vuong study the independent private value *single object* first price auction. Adapting it to a procurement model, their bidders solve the following optimization problem:

$$\pi^i(c^i) = \max_{b^i \in \mathbb{R}} (b^i - c^i) G(b^i)$$

where $G(.)$ is bidder i's probability of winning as a function of his bid. Function G is monotonically decreasing, so this optimization problem is also a.e. differentiable with first order condition:

$$(b^i - c^i)G'(b^i) = -G(b^i) \tag{4}$$

Guerre, Perrigne and Vuong appeal to the known characterization of the equilibrium in the independent private values first price auction to argue that $G'(b^i)$ is strictly positive everywhere. Thus, we can rewrite (4) as

$$c^i = b^i + \frac{G(b^i)}{G'(b^i)} \tag{5}$$

In words: when the bidders' optimization problem is differentiable, its first order condition always identifies costs (given b^i, there is a unique value for c^i that solves (5)). Thus $\mathcal{B}_{i2} = \emptyset$. By contrast, we will see below that $\mathcal{B}_{i2}$ is not necessarily empty in the combinatorial first price auction. In addition, equilibrium strategies in the independent private values single object first price auction are strictly increasing . Hence $\mu(\mathcal{B}_{i3}) = 0$ (which follows from the fact that the bidder's optimization problem is a.e. differentiable) implies $\eta^i(\mathcal{B}_{i3}) = 0$.

Putting both elements together yields $\eta^i(\mathcal{B}_{i1}) = 1$: the model is identified on the basis of the first order conditions. But notice that, on the way, Guerre, Perrigne and Vuong have relied on some knowledge about the nature of equilibrium in the auction.

The next Lemma provides the needed equilibrium characterization results for the combinatorial first price auction. In particular, it establishes that $\eta^i(\mathcal{B}_{i3}) = 0$ too in the combinatorial first price auction.

Lemma 1 *Consider the private value combinatorial first auction model described in Section 2, without a reserve price or with a known reserve price, and with or without combination bid constraints. Then, in any equilibrium where bidders bid on or above their costs:*
(1) Bidders are indifferent about the way in which ties that occur with positive probability are resolved,
(2) $\lim_{\mathbf{b}\to\mathbf{b}^*} \sum_{s\subseteq S}(b_s - c^i_s)G_s(\mathbf{b}) = \sum_{s\subseteq S}(b^*_s - c^i_s)G_s(\mathbf{b}^*)$ *for all* $\mathbf{b}^* \in \mathcal{B}_i$*, such that* $\mathbf{b}^*$ *is bidder i's optimal bid given costs* $\mathbf{c}^i$.
(3) $\eta^i(\mathcal{B}_{i3}) = 0$ *for all i.*

The proof of Lemma 1 can be found in the Appendix. Part 1 of Lemma 1 ensures that an equilibrium exists with the London bus routes tie-breaking rule. Part 2 of Lemma 1 says that bidders' expected equilibrium payoff is continuous in their bids, at their equilibrium bids. A (locally) monotone and continuous function is always left and right differentiable. Thus, bidders' expected equilibrium payoffs are left and right differentiable at their equilibrium bids. Probabilities of winning are in principle not left and right differentiable at discontinuity points. However, because bidders are indifferent among the ways in which those ties are resolved, we can adopt without loss of generality the convention that one-sided derivatives at such points equal their limits. Formally, suppose $\mathbf{b}^i$ corresponds to a discontinuity point of G_s with respect to b_t. Then its left derivative with respect to b_t at $\mathbf{b}^i$ is defined as $\frac{\partial}{\partial b_t}G_s(\mathbf{b}^i) = \lim_{\widehat{\mathbf{b}}^i \uparrow \mathbf{b}^i} \lim_{\varepsilon \downarrow 0} \frac{G_s(\widehat{b}^i_t - \varepsilon, \widehat{\mathbf{b}}^i_{-t}) - G_s(\widehat{\mathbf{b}}^i)}{\varepsilon}$. Finally, part 3 establishes that there are no mass point in the distribution of equilibrium bids at kinks in the objective function.[10] As a consequence, we can ignore bids in $\mathcal{B}_{i3}$ for identification purposes because they are submitted by a zero measure of cost realizations.

3.2 Identification conditions

At any point where bidder i's objective function is differentiable, his optimal bid vector must satisfy the first order conditions:[11]

$$G_t(\mathbf{b}^i) + \sum_{s \subseteq S} (b^i_s - c^i_s) \frac{\partial}{\partial b_t} G_s(\mathbf{b}^i) = 0 \qquad t \subseteq S$$

or, in matrix notation:

$$\nabla \mathbf{G}(\mathbf{b}^i)[\mathbf{b}^i - \mathbf{c}^i] = -\mathbf{G}(\mathbf{b}^i) \tag{6}$$

where the $(2^m - 1)$ by $(2^m - 1)$ matrix $\nabla \mathbf{G}(\mathbf{b}^i)$ is defined by $\nabla G_{t,s}(\mathbf{b}^i) = \frac{\partial}{\partial b_t} G_s(\mathbf{b}^i)$ for $s, t \subseteq S$ and $\mathbf{G}(\mathbf{b}^i)$ is a $2^m - 1$ x 1 vector with $G_s(\mathbf{b}^i)$ as components.

The first order conditions define a system of linear equations in the unknown costs, $\mathbf{c}^i$. Identification of costs on the basis of the first order conditions then reduces to the question of existence and uniqueness of a solution to this system.

Lemma 2 (Sufficient condition for identification) *A sufficient condition for identification in the combinatorial first price auction is that $\nabla \mathbf{G}(\mathbf{b}^i)$ is invertible for all i and all equilibrium bids $\mathbf{b}^i$.*

Proof. The first order conditions in (6) define a system of linear equations in the unknown cost parameters (the $[\mathbf{b}^i - \mathbf{c}^i]$ vector). The invertibility of matrix $\nabla \mathbf{G}(\mathbf{b}^i)$ is a necessary and sufficient condition for a unique solution $\mathbf{c}^i = \boldsymbol{\phi}^i(\mathbf{b}^i) \in \mathbb{R}^{2^m - 1}$. Thus $\mathcal{B}_{i2}$ is

[10] The definition of mass points in multi-dimensions is a straightforward extension of their definition for univariate distribution: $\eta^i(.)$ is said to have a mass point at $B \subseteq \mathcal{B}_i$ if $\eta^i(B) > 0$ but $\mu(B) = 0$.

[11] This is independent of whether the equilibrium is in pure or mixed strategies. Even if the equilibrium is in mixed strategies, equilibrium bids must satisfy the first order conditions for optimal bids.

empty. Because $\eta^i(\mathcal{B}_{i3}) = 0$ by Lemma 1, identification of the distribution of costs F follows directly. ■

Note that Lemma 2 says that point identification of the cost vectors is a sufficient, but not a necessary, condition for the identification of the distribution of costs. Indeed, under some restrictive conditions like independence, multidimensional distributions can be identified from the observation of lower dimensional data (Berman, 1963). While the equilibrium bids submitted by a bidder on the different contracts are unlikely to be independently distributed, there is no known result confirming this.

The next lemma (proved in the Appendix) investigates the properties of matrix $\mathbf{\nabla G}$:

Lemma 3 (Properties of $\triangledown$G) *Consider matrix* $\mathbf{\nabla G}(\mathbf{b}^i)$ *(with elements defined by* $\nabla G_{t,s}(\mathbf{b}^i) = \frac{\partial}{\partial b_t} G_s(\mathbf{b}^i)$*) evaluated at any optimal bid vector* $\mathbf{b}^i$ *by bidder* i*. Then:*

(1) $\frac{\partial}{\partial b_t} G_t(\mathbf{b}^i) \leq 0$ *for all* t*, and strictly so if* $G_t(\mathbf{b}^i) > 0$*.*

(2) $\frac{\partial}{\partial b_t} G_s(\mathbf{b}^i) \geq 0$ *for all* $t \neq s$*.*

(3) $\sum_{s \subseteq S} \frac{\partial}{\partial b_t} G_s(\mathbf{b}^i) \leq 0$ *for all* t*, and strictly so for some* t *if there exists* $s \subseteq S$ *such that* $G_s(\mathbf{b}^i) > 0$*.*

(4) $\mathbf{\nabla G}(\mathbf{b}^i)$ *is invertible if* $G_s(\mathbf{b}^i) > 0$ *for all* s*.*

(5) The determinant of any submatrix made from removing some rows and the corresponding columns of $\mathbf{\nabla G}$ *has sign (-1)*^{r} *where r is the number of remaining rows/columns if all bid components in* $\mathbf{b}^i$ *except those removed have a strictly positive probability of winning.*

Note: Lemma 3 also applies to the Jacobian matrix of the probabilities of winning for the optimization problem where bidder i is constrained to bid only on a subset $K \subset 2^S$ of the contracts, instead of all $2^m - 1$ contracts.

The first two properties say that the probability of winning a given route cannot increase when the bidder increases his bid on that route, while the probability of winning any other route cannot decrease. Combining these two properties implies that, when a bidder wins a route with strictly positive probability, increasing his bid on this route must strictly decrease his probability of winning that route. Otherwise, increasing his bid on that route would constitute a profitable deviation, violating the assumption that $\mathbf{b}^i$ is an optimal bid.

Property 3 says that when a bidder increases one of his bids, his probability of winning anything cannot increase. Moreover, if he is winning at least one route with strictly positive probability at $\mathbf{b}^i$, his probability of winning anything must be strictly decreasing. (Again, otherwise, he would have a profitable deviation.)

Finally, property 4 relies on properties 1, 2 and 3, and on the property shown in the Appendix that, at the optimum, bids with a strictly positive probability of winning compete with one another.

From Lemma 3, $\nabla \mathbf{G}$ is invertible when all bids have a strictly positive probability of winning. However, not all equilibrium bids need to satisfy this condition. Irrelevant bids (bids that never win) can be optimal from a bidder's perspective because of the strategic price discrimination motivation for combination bids: submitting a bid that never wins on a contract ensures that this bid does not compete with a potentially more profitable bid.[12] Such bids are problematic for identification. Formally, bid b_s^i is irrelevant if $G_s(b_s^i, \mathbf{b}_{-s}^i)$ $= 0$ and there exists $\varepsilon > 0$ such that $G_s(b_s^i - \varepsilon, \mathbf{b}_{-s}^i) = 0$ ($\mathbf{b}_{-s}^i$ denotes bidder i's bids on the other contracts but s). Suppose bidder i submitted an irrelevant bid on contract s. Then, any alternative bid vector $(\widehat{b}_s^i, \mathbf{b}_{-s}^i)$ with $\widehat{b}_s^i > b_s^i$ would have been equally optimal for bidder i, and therefore equally informative. More formally, $\frac{\partial}{\partial b_s} G_t(\mathbf{b}^i) = 0$ (small changes in b_s^i do not affect the probability that bidder i wins contract s or any other contracts) and $\frac{\partial}{\partial b_t} G_s(\mathbf{b}^i) = 0$ for all t. The row and column corresponding to contract s in matrix $\nabla \mathbf{G}$ are all zeros. Therefore $\nabla \mathbf{G}$ cannot be inverted and the condition in Lemma 2 fails.

To investigate identification in the presence of irrelevant bids, suppose that, at equilibrium, bidder i submitted a bid with a strictly positive probability of winning on all contracts in $K \subset 2^S$. Let $\mathbf{b}_K^i$ denote his bid vector on contracts in K and define by $\mathbf{c}_K^i$ the corresponding subvector of costs. Consider the following alternative optimization problem:

$$\widehat{\pi}^i(\mathbf{c}^i) = \max_{\mathbf{b}_K^i} \sum_{s \in K} (b_s^i - c_s^i) \widehat{G}_s(\mathbf{b}_K^i) \tag{P2}$$

where $\widehat{G}_s(.) = G_s(., \; b_t^i = \infty, t \notin K)$. Because (P2) is a constrained version of (P1), $\widehat{\pi}^i(\mathbf{c}^i) \leq \pi^i(\mathbf{c}^i)$ for all $\mathbf{c}^i$. The two expressions are equal if the optimal bid in (P1) is such that bidder i only wins contracts in K with strictly positive probability. When this is the case, the solutions to the two problems coincide and the first order conditions of (P2) describe equally well the optimal bids:

$$\widehat{G}_w(\mathbf{b}_K^i) + \sum_{s \in K} (b_s^i - c_s) \frac{\partial}{\partial b_w} \widehat{G}_s(\mathbf{b}_K^i) = 0 \qquad w \in K \tag{7}$$

This $|K|$ dimensional system of linear equations can again be written in matrix notation: $\nabla \widehat{\mathbf{G}}(\mathbf{b}_K^i)[\mathbf{b}_K^i - \mathbf{c}_K^i] = -\widehat{\mathbf{G}}(\mathbf{b}_K^i)$. By Lemma 3, matrix $\nabla \widehat{\mathbf{G}}(\mathbf{b}_K^i)$ is invertible. Therefore, costs $\mathbf{c}_K^i$ are identified:

$$\mathbf{c}_K^i = \mathbf{b}_K^i + \nabla \widehat{\mathbf{G}}(\mathbf{b}_K^i)^{-1} \widehat{\mathbf{G}}(\mathbf{b}_K^i) \tag{8}$$

[12] Armstrong (1996) provides a decision-theoretic example where this property holds at the optimum.

Irrelevant bids do not affect the identification of the costs of the other bids. This is somewhat remarkable in this multi-unit auction setting where costs are *a priori* jointly determined as the solution to a system of equations. The reason is that, in the case of irrelevant bids, bids on *other* contracts do not affect the probability that bidder i wins the contract on which he submitted an irrelevant bid (it remains zero). Likewise, irrelevant bids do not affect the probability of winning any of the other contracts. This removes the interdependency among the first order conditions.

Now consider $s \notin K$. Let $\overline{b}_s^i(\mathbf{b}_{-s}^i)$ be the threshold for b_s^i over which it becomes irrelevant that is:

$$\overline{b}_s^i(\mathbf{b}_{-s}^i) = \inf\{b_s^i : b_s^i \text{ is irrelevant given } \mathbf{b}_{-s}^i\}$$

($\overline{b}_s^i(\mathbf{b}_{-s}^i)$ is a function of bidder i's bids on the other contracts as well as the support of competitors' bids.) Since expected profit is continuous, it must be that the slope of bidder i's expected profit to the left of $\overline{b}_s^i(\mathbf{b}_{-s}^i)$ is positive. Formally,

$$\lim_{b_s \uparrow \overline{b}_s^i} G_s(b_s, \mathbf{b}_{-s}^i) + \sum_{t \in K} (b_t^i - c_t^i) \frac{\partial}{\partial b_s} G_t(\overline{b}_s^i, \mathbf{b}_{-s}^i) + (\overline{b}_s^i - c_s^i) \frac{\partial}{\partial b_s} G_s(\overline{b}_s^i, \mathbf{b}_{-s}^i) \geq 0 \tag{9}$$

where the derivatives are all left derivatives.[13,14] Given (8), c_t^i, $t \in K$, are point identified. Thus, the only unknown in (9) is c_s^i. Because $\frac{\partial}{\partial b_s} G_s < 0$, (9) identifies a lower bound to the cost c_s^i, $\underline{c}_s^i$. Repeating this procedure for every irrelevant bid in $\mathbf{b}^i$ and combining with (8) for the other bids, identifies a lower bound, $\underline{\mathbf{c}}^i$, to the true cost vector $\mathbf{c}^i : \underline{\mathbf{c}}^i \leq \mathbf{c}^i$ (in the component-wise order). Let $\underline{\phi}^i(.)$ the function that associates bidder i's equilibrium bid vectors to the lower bound cost vector that rationalizes them, i.e. $\underline{\mathbf{c}}^i = \underline{\phi}(\mathbf{b}^i)$. We have $F_{iH}(\mathbf{c}) = \eta^i(\{\mathbf{b}^i \in \mathcal{B}_{i1} \cup \mathcal{B}_{i2} : \underline{\phi}(\mathbf{b}^i) \leq \mathbf{c}\})$.

Likewise, we can define an upper bound on the costs that rationalize the observed bids. Let $\overline{c}_s^i = c_s^i$ if bidder i's cost on contract s is point identified. Let $\overline{c}_s^i = \infty$ otherwise. Let $\overline{\phi}^i(\mathbf{b}^i) = \overline{\mathbf{c}}^i$, and define $F_{iL}(\mathbf{c}) = \eta^i(\{\mathbf{b}^i \in \mathcal{B}_{i1} \cup \mathcal{B}_{i2} : \overline{\phi}(\mathbf{b}^i) \leq \mathbf{c}\})$.

The next Theorem summarizes the results thus far:

Theorem 1 (Identification and partial identification) *(i) Consider any bid* $\mathbf{b}^i \in \mathcal{B}_{i1} \cup \mathcal{B}_{i2}$. *If* $G_s(\mathbf{b}^i) > 0$, *then* c_s^i *is point identified. Otherwise,* $c_s^i \geq \underline{c}^i{}_s$, *where* $\underline{c}^i{}_s$ *is defined by (9).*

[13] In (9), the sum is only over $t \in K$ because $\frac{\partial}{\partial b_s} G_w(\mathbf{b}^i) = 0$ for any $w \notin K$ and $w \neq s$.

[14] The first term in (9) takes into account that G_s may be discontinuous at $\overline{b}_s^i$. In that case, see the discussion after Lemma 1 for the construction of the left derivatives.

(ii) If all equilibrium bids have a strictly positive probability of winning, the combinatorial first price model is identified. If equilibrium strategies contain bids that have zero probability of winning, the model is partially identified. Any F_i such that $F_{iH}(\mathbf{c}) \geq F_i(\mathbf{c}) \geq F_{iL}(\mathbf{c})$ for all $\mathbf{c}$ (F_{iH} and F_{iL} defined in the text) is consistent with the observed equilibrium.

Proof. We first consider the identification of the set of costs associated with each equilibrium bid $\mathbf{b}^i \in \mathcal{B}_{i1} \cup \mathcal{B}_{i2}$. Suppose that $G_s(\mathbf{b}^i) > 0$ for all s. From Lemma 3(4), the Jacobian matix $\nabla \mathbf{G}$ is invertible, and Lemma 2 implies that the costs rationalizing $\mathbf{b}^i$ are uniquely determined. Suppose now that $G_s(\mathbf{b}^i) = 0$ for some s which means that $\mathbf{b}^i$ includes an irrelevant bid. Irrelevant bids correspond to bids in $\mathcal{B}_{i2}$ for which the arguments around (8) and (9) establish bounds on their associated costs. This establishes (i) and pins down the correspondence $\phi^i : \mathcal{B}_{i1} \cup \mathcal{B}_{i2} \rightrightarrows \mathbb{R}^{2^m-1}$.
We now turn to the question of the identification of F_i. By Lemma 1(3), we can without loss of generality focus on bids in $\mathcal{B}_{i1} \cup \mathcal{B}_{i2}$ because bids in $\mathcal{B}_{i3}$ are submitted by a zero measure of costs. When $G_s(\mathbf{b}^i) > 0$ for all s, for all $\mathbf{b}^i$ in $\mathcal{B}_{i1} \cup \mathcal{B}_{i2}$ and for all i, ϕ^i is a function for all i and identification of F_i follows from Lemma 2. Otherwise, two selections of ϕ^i of interest are its minimum selection, $\underline{\phi}^i(.)$, and its maximum selection, $\overline{\phi}^i(.)$. They generate F_{iH} and F_{iL} defined as $F_{iH}(\mathbf{c}) = \eta^i(\{\mathbf{b}^i \in \mathcal{B}_{i1} \cup \mathcal{B}_{i2} : \underline{\phi}(\mathbf{b}^i) \leq \mathbf{c}\})$ and $F_{iL}(\mathbf{c}) = \eta^i(\{\mathbf{b}^i \in \mathcal{B}_{i1} \cup \mathcal{B}_{i2} : \overline{\phi}(\mathbf{b}^i) \leq \mathbf{c}\})$ for $\mathbf{c} \in \mathbb{R}^{2^m-1}$. Because cost ranges are identified, any $F_i \in \mathcal{F}$ such that $F_{iH} \geq F_i \geq F_{iL}$ is consistent with the observed equilibrium. ■

Theorem 1 provides a partial identification result when equilibrium behavior involves irrelevant bids. Given the results in Armstrong (1996), this is likely to be a generic case. Nevertheless, partial identification results can be used for answering policy questions by estimating policy variables for all cost distributions in the set of cost distributions consistent with equilibrium behavior. Moreover, bounds on policy variables of interest are often all we are interested in. The bounds derived in Theorem 1 allow us to compute bounds on the expected procurement cost in a Vickrey auction or in any other auction where bid functions are monotonic in costs.[15] Note that the bounds on the distribution of costs are essentially sharp in the sense that F_{iH} and F_{iL} correspond - modulo a slight pertubation to ensure a continuous density - to feasible cost distributions that is consistent with the observed equilibrium bids.

[15] The condition needed here is that the policy variable be monotonic in some (partial) order of the elements of the set. See Manski and Tamer (2002) for related results in the presence of partial identification. For another example where the bounds on primitives are directly useful to compute bounds on policy variables in auctions, see Haile and Tamer (2003).

An interpretation of Theorem 1 is that identification requires the dimensionality of the observed information to match that of the information to infer. The set of bids that have a strictly positive probability of winning determines the dimensionality of the observed information. Any irrelevant bid introduces a dimension of under-identification, and it is at the source of the partial identification result of Theorem 1. Pursuing on this intuition, we can easily prove the following Corollary:

Corollary 1 *The multi-unit first price auction model is not identified if only stand-alone bids are permitted (i.e.* $\mathbf{b}^i \in \mathbb{R}^m$*).*
The multi-unit first price auction model is identified if the following two conditions hold:
(i) costs are additive, i.e. $c_{s \cup s'} = c_s + c_{s'}$ *for all* $s, s' \subset S$ *with* $s \cap s' = \varnothing$*.*
(ii) all stand-alone bids have a strictly positive probability of winning, i.e., $G_s(\mathbf{b}^i) > 0$ *for* $s \in S$
These two conditions will be satisfied if bidders' supports of cost realizations have the same upper bound and there is no reserve price.

The intuition is straightforward. In the multi-unit auction model, the underlying private information to infer (the costs c_s) is $2^m - 1$ dimensional. When only stand-alone bids are permitted, the observed information is at most m dimensional. So, there is no hope to infer costs, unless the dimensionality of private information is also m (note that additive costs are ruled out by assumption 1).

Proof. When bids are only permitted on individual routes, the auction effectively become m independent single object first price auctions. The necessary conditions for optimal bids are the m first order conditions. This is a system of m linear equations in $2^m - 1$ variables (the unobserved c^i_w). This system is under-identified. Thus costs are not identified and as a result the model is not identified.
When costs are additive, identification follows from the arguments in Guerre, Perrigne and Vuong (2000) Campo, Perrigne and Vuong (2003) and Athey and Haile (2005) applied to the m independent auctions. The conditions of same upper bound to cost realizations and no reserve price guarantee all equilibrium bids have a strictly positive probability of winning and the model is identified through the first order conditions. ■

3.3 Other constraints on bidding behavior

Most real-life multi-unit auctions include various restrictions on the set of allowable bids which may further reduce the dimensionality of the observed information. In this section, we show how to extend our identification results and derive identification bounds on the costs and distributions of costs in these cases. We illustrate our approach by considering two types of restrictions present in our data. First, the rule of the auction imposes that bids on a combination of routes must be no greater than the sum of the constituent bids. Second, London Transport Buses imposes a reserve price. Bidders are not obliged to submit bids on all routes and some bidders indeed submit bids only on a subset of the routes auctioned. Our interpretation is that it was not profitable for these bidders to submit a bid that would have had a positive chance of winning.

We first consider the combination bid constraint. In the presence of combination bid constraints, the bidders' optimization problem becomes:

$$\max_{\mathbf{b}^i} \sum_{s \subseteq S} (b_s^i - c_s) G_s(\mathbf{b}^i)$$

subject to:

$$b_s^i \leq b_w^i + b_t^i \text{ for all } s, t, w \subseteq S \text{ such that } t \cap w = \emptyset \text{ and } t \cup w = s$$

This optimization is differentiable almost everywhere (and continuous at all optimal bids from Lemma 1). Its first order conditions are given by

$$G_t(\mathbf{b}^i) + \sum_{s \subseteq S} (b_s^i - c_s^i) G_s^t(\mathbf{b}^i) - \sum_{r,w \subseteq S} \lambda_{t=r\cup w} + \sum_{s,r \subseteq S} \lambda_{s=r\cup t} = 0 \qquad \forall t \subseteq S$$

where $\lambda_{t=r\cup w}$ is the multiplier of the constraint $b_t^i \leq b_r^i + b_w^i$ ($\lambda_{t=r\cup w} \geq 0$). We define a new column vector $\mathbf{D}(\mathbf{b}^i, \boldsymbol{\lambda})$ with component $D_t(\mathbf{b}^i, \boldsymbol{\lambda}) = -G_t(\mathbf{b}^i) - \sum_{s,r} \lambda_{s=r\cup t} + \sum_{r,w} \lambda_{t=r\cup w}$. The first order conditions can then be rewritten as

$$\boldsymbol{\nabla}\mathbf{G}(\mathbf{b}^i)(\mathbf{b}^i - \mathbf{c}^i) = \mathbf{D}(\mathbf{b}^i, \boldsymbol{\lambda})$$

Lemma 3 applies to matrix $\boldsymbol{\nabla}\mathbf{G}$ (with the convention that where a combination bid constraint binds, the one-sided derivative that satisfies the constraint is used). If all bids have a strictly positive probability of winning, $\boldsymbol{\nabla}\mathbf{G}$ can be inverted and $\mathbf{c}^i$ solves

$$\mathbf{c}^i = \mathbf{b}^i - \boldsymbol{\nabla}\mathbf{G}(\mathbf{b}^i)^{-1}\mathbf{D}(\mathbf{b}^i, \boldsymbol{\lambda}) \tag{10}$$

If a subset of the equilibrium bids are irrelevant, $\boldsymbol{\nabla}\mathbf{G}$ cannot be inverted. In that case, we can proceed as before by considering the alternative optimization problem where bidder i only bids on the contracts $K \subset 2^S$ for which he submitted a bid with a strictly positive probability of winning in the original problem:

$$\widehat{\pi}^i(\mathbf{c}^i) = \max_{\mathbf{b}^i_K} \sum_{s\in K} (b^i_s - c^i_s)\widehat{G}_s(\mathbf{b}^i_K)$$

$$\text{s.t. } b^i_s \quad \leq \quad b^i_w + b^i_t \text{ for all } s,t,w \in K \text{ such that } t\cap w = \emptyset \text{ and } t\cup w = s$$

The associated $|K|$ by $|K|$ matrix $\boldsymbol{\nabla}\widehat{\mathbf{G}}_s(\mathbf{b}^i_K)$ satisfies the conditions of Lemma 3(4) and is therefore invertible, yielding:

$$\mathbf{c}^i_K = \mathbf{b}^i_K - \boldsymbol{\nabla}\widehat{\mathbf{G}}(\mathbf{b}^i_K)^{-1}\mathbf{D}(\mathbf{b}^i_K, \boldsymbol{\lambda}) \tag{11}$$

A direct consequence of expressions (10) and (11) is that the costs of the contracts which bidder i wins with strictly positive probability at equilibrium are uniquely identified up to the value of the multiplier $\boldsymbol{\lambda}$. If no constraint binds so that all multipliers are equal to zero, the costs are point identified. When any of the constraints binds, bounds on each cost can be constructed from (11) by exploiting the fact that the multipliers on the binding constraints are all positive. Formally, let

$$\overline{c}^i_s \quad = \quad \sup_{\boldsymbol{\lambda}\geq 0}\{\text{component } s \text{ of vector } \mathbf{b}^i_K - \boldsymbol{\nabla}\widehat{\mathbf{G}}(\mathbf{b}^i_K)^{-1}\mathbf{D}(\mathbf{b}^i_K, \boldsymbol{\lambda})\} \tag{12}$$

$$\underline{c}^i_s \quad = \quad \inf_{\boldsymbol{\lambda}\geq \mathbf{0}}\{\text{component } s \text{ of vector } \mathbf{b}^i_K - \boldsymbol{\nabla}\widehat{\mathbf{G}}(\mathbf{b}^i_K)^{-1}\mathbf{D}(\mathbf{b}^i_K, \boldsymbol{\lambda})\} \tag{13}$$

We thus have: $\underline{\mathbf{c}}^i_K \leq \mathbf{c}^i_K \leq \overline{\mathbf{c}}^i_K$.

We now turn to reserve prices. When the auction includes reserve prices but no combination bid constraint, the formal treatment of non submitted bids is exactly the same as that of irrelevant bids, except for the fact that the threshold for irrelevant bids is replaced by:

$$\overline{b}^i_s(\mathbf{b}^i_{-s}) = \min\{R_s, \inf\{b^i_s : b^i_s \text{ is irrelevant given } \mathbf{b}^i_{-s}\}\}$$

where R_s is the known reserve price on s. Costs associated with bids that have a strictly positive probability of winning are point identified following (8). Costs associated with irrelevant bids or non submitted bids are partially identified based on (9) using the new definition of $\overline{b}^i_s(\mathbf{b}^i_{-s})$ above.

When the auction includes a reserve price and a combination bid constraint, a binding reserve price does not always identify a lower bound to the associated cost. To get some intuition for this result, consider the following two-route auction. Suppose bidder i only submitted a bid on route 1. Using the earlier results, we can point identify c^i_1 and place a lower bound on c_{12}. However, bidder i's bidding behavior does not place a lower bound on c_2 if $b_1^i + R_2 < R_{12}$. By not submitting a bid that has a positive probability of winning on route 2, bidder i also ensures that he will never win both routes. This could be a best response if c_{12} is very large.

Theorem 2 summarizes what can be identified in the presence of combination bid constraints and reserve prices.

Theorem 2 (Partial identification with reserve prices and/or combination bid constraints) *Consider the combinatorial first price auction model with reserve prices and, possibly, a combination bid constraint.*

*(1) **Identification of costs.** Consider any equilibrium bid vector $\mathbf{b}^i \in \mathcal{B}_{i1} \cup \mathcal{B}_{i2}$.*

(i) Suppose $G_s(\mathbf{b}^i) > 0$ for some $s \subseteq S$, then upper and lower bounds to c_s^i are identified following (12) and (13). In particular, if $G_s(\mathbf{b}^i) > 0$ for all $s \subseteq S$ and $\mathbf{b}^i$ satisfies strictly all combination bid constraints (if any applies), then $\mathbf{c}^i$ is point identified.

(ii) Suppose $G_s(\mathbf{b}^i) = 0$ for some $s \subseteq S$ and the auction does not have a combination bid constraint. Then a lower bound to c_s^i, $\underline{c}_s^i$, is identified following (9). If $\overline{b}_s^i(b_{-s}^i) = R_s$ then $\underline{c}_s^i = R_s$.

(iii) Suppose $G_s(\mathbf{b}^i) = 0$ for some $s \subseteq S$ and the auction has a combination bid constraint. Then a lower bound to c_s^i, $\underline{c}_s^i$, is identified following (9) if $G_t(\mathbf{b}^i) > 0$ or $G_t(\overline{b}_s^i(\mathbf{b}_{-s}^i), \mathbf{b}_{-s}^i) = 0$ for all $t \supset s$. No bound is identified otherwise.

*(2) **(Partial) Identification of the distributions of costs.** For all i, let $\phi^i(.) : \mathcal{B}_{i1} \cup \mathcal{B}_{i2} \rightrightarrows \mathbb{R}^{2^m-1}$ define the mapping from bidder i's equilibrium bids to the set of costs (obtained from the application of (1)) that are consistent with them. Any $F_i \in \mathcal{F}$ such that $F^i(\mathbf{c}) = \eta^i(\{\mathbf{b}^i \in \mathcal{B}_{i1} \cup \mathcal{B}_{i2} : \widetilde{\phi}^i(\mathbf{b}^i) \leq \mathbf{c}\})$ for some selection $\widetilde{\phi}^i$ of ϕ^i, is consistent with the observed equilibrium.*

Part (1) of Theorem 2 follows from the arguments in the text above, except for the claim that $\underline{c}_s^i = R_s$, which is proved in the Appendix. Part (1) identifies costs or cost ranges for bids that correspond to $\mathcal{B}_{i1}$ and $\mathcal{B}_{i2}$. Part (2) follows from part (1) given that bids in $\mathcal{B}_{i3}$ are submitted by a zero measure of costs.

In some cases, the value of $\boldsymbol{\lambda}$ that solves for the bounds in (12) and (13) can be known *a priori*. The next Theorem characterizes this value of $\boldsymbol{\lambda}$.

Theorem 3 *Consider any optimal bid $\mathbf{b}^i$:*
(1) For any s such that $G_s(\mathbf{b}^i) > 0$, c_s^i depends positively on the value of $\lambda_{s=t\cup w}$, and negatively on $\lambda_{t=s\cup w}$ for all t and w.

(2) Consider any t, $w \subseteq S$ such that $t \cap w = \emptyset$. If the combination bid constraint for these contracts is the only binding combination bid constraint at $\mathbf{b}^i$, an upper bound to the synergy involved between these two contracts is given by the solution $c_t + c_w - c_{t \cup w}$ of the system in (11) when the Lagrangian multiplier $\lambda_{\{t \cup w\}=t \cup w} = 0$.

Theorem 3 uses Cramer's rule and the properties of the determinants to sign how the solution in (10) and (11) depends on the value of the multiplier. Theorem 3 is silent concerning how c_s^i depends on $\lambda_{r=t \cup w}$ for $r, t, w \neq s$. This relationship cannot be signed on a systematic basis. Instead, (12) and (13) should be used.

Note that, in principle, we could have generated a lower and upper bound F_{iL} and F_{iH} to the distribution consistent with the observed bid distribution. However, a consequence of Theorem 3 is that such F_{iL} and F_{iH} would never be sharp in the presence of a combination bid constraint, unlike the bounds on the distributions of costs derived in Theorem 1. Because the bounds on costs for different contracts correspond to different values of the multipliers, they do not correspond to a cost realization consistent with the observed bid. Thus, the upper and lower bounds on the cost distributions do not rationalize the observed bid distribution. As a result, they are less useful for policy analysis. The whole set of identified distributions rather than the bounds themselves should be used for such purposes.

4 The London Bus Market

This section describes the London bus market, gives descriptive summaries of our data and motivates the empirical specification described in section 6.

The London bus market represents about 800 routes serving an area of 1,630 square kilometers and more than 3.5 million passengers per day. It is valued at 600 million Pounds per year (US $1,100 million). Deregulation was introduced by the London Regional Transport Act of 1984. The Transport Act designated London Regional Transport (LRT) as the authority responsible for the provision and procurement of public transport services in the Greater London area, as well as the development and operations of bus stations and the network-wide operational maintenance. Private procurement was encouraged. In order to enhance competition, LRT, which by virtue of the Transport Act acted as the holding company for the original public operator London Buses Limited, created a separate tendering division, independent from its operational division, and split the formerly unitary London Buses into 12 operational subsidiairies. These were privatized in 1994. In practice, the introduction of route tendering was very gradual. The first tenders took place in 1985, but it was not until 1995 that half of the network was tendered at least once.[16] Since then, tendering has reached its steady state regime with 15-20% of the network tendered every year.

The procurement process. About every two weeks London Transport Buses issues an invitation to tender which provides a detailed description of upcoming contracts for sale. The invitation simultaneously covers several routes, usually in the same area of London. For each route, the invitation provides a complete description of the service for tender including the routing, service frequency and vehicle type. Contract length is typically five years. A set of pre-qualified bidders may submit sealed bids for individual routes. In addition, bidders may submit a bid for route combinations within the auction. A bid specifies an annual price at which the bidder is willing to provide the service.[17] There is a period of two months between the invitation to tender and the tender return date, and another two months before contracts are awarded. The official award criterion is *best economic value* and the process follows EU law for fair competition. In practice, this means that the contract is awarded to the low bidder but deviations at the margin are possible to account for bidder quality for instance.[18] To allow winning bidders to reorganize and order new buses if necessary, contracts start 8 to 10 months after the award date.

Description of the bid data.[19] We have collected data on 179 auctions consisting of a total of 674 routes offered to bidders between December 1995 and May 2001 (return date). For each auction and for each route in the auction, the data include the following information: (1) contract duration and planned start of the contract (2) route characteristics including the route start and end points; route type (day route, night route, school service,

[16] Non-tendered routes remained operated by the subsidiaries of London Buses Limited under a negotiated block grant. The private operators and the subsidiaries competed for the tendered services.

[17] London Transport Buses has experimented with different contractual forms. The majority of contracts are so called gross cost contracts, in which the revenues collected on the buses accrue to London Transport Buses and the operator receives a fixed fee for the service. Some contracts are net cost contracts, in which the operators take responsibility for the revenues. The price for the operator service then consists of those revenues plus a transfer from (or payment to) London Transport Buses. Finally, net cost contracts may contain a provision that limits the risk the operator takes in case the revenues were too different from the forecast. If bidders are risk neutral, which we assume in our analysis, all three contracts forms are equivalent.

[18] From the 118 auctions studied in the empirical analysis, a deviation at the margin for some route occurs on 39 auctions. The empirical analysis revealed no systematic patterns in these considerations that we could model explicitly. We interpret the considerations at the margin as noise in the awarding process.

[19] Appendix A provides further details concerning the sources of the data.

mobility route); annual mileage; bus type (single deck, midibuses, double deck or routemaster); and the peak vehicle requirement;[20] (3) the identity of bidders and all their submitted bids (including bids for combinations of the routes in the auction). For the auctions held starting in May 2000, the data also contain an internal cost estimate generated by London Transport Buses for every route. All price data are expressed in December 1995 Pounds.

Contract heterogeneity. There are many dimensions along which the routes in our sample vary. Route characteristics affect costs and, ultimately, participation and bids. A monetary measure of contract heterogeneity is the internal cost estimate (ICE) prepared by London Transport Buses since May 2000. We generated a predicted internal cost estimate based on a regression of the ICE on route characteristics.[21] We found the predicted ICE to be an accurate assessment of the final cost. We considered a regression of the log of bids and the log of low bids on the log of the internal cost estimate. The log internal cost estimate explains 93% of the variation in the log bids. In order keep the number of explanatory variables in our empirical specification small, we use the predicted internal cost estimate to account for contract heterogeneity.

Most auctions consist of only few routes.[22] Our estimation uses the 118 auctions in our data that have no more than 3 routes. Table 1 provides summary statistics of our bid data for these auctions.[23]

Table 1 Descriptive Summary of the Bid Data (Auctions with 1, 2 and 3 routes)

Variable	Obs	Mean	Std	Min	Max
ln(ICE)	218	13.28	1.29	10.82	15.56
Number-of-Actual-Bidders-per-Auction	118	3.70	1.74	1	8
Number-of-Actual-Bidders-per-Route	218	2.94	1.57	1	7
Log-Stand-Alone-Bid	641	13.12	1.28	9.47	15.87
Log-Combination-Bid	83	14.48	0.74	11.75	15.89
Money-Left-on-Table (%)	177	13.54	20.36	0.06	157.86

[20] The peak vehicle requirement determines how many buses the winning operator needs to commit to the contract.

[21] The route characteristics include the peak vehicle requirement, annual mileage for central London routes and annual mileage for non-central London routes interacted with bus type dummies (single deck or midibus, double deck or route master).

[22] The distribution of routes across tranches in our sample is the following: 50 tranches consist of a single route, 36 tranches have two routes, 32 tranches have 3 routes, 13 tranches have 4 routes, 10 have 5 routes, 27 tranches have between 6 and 10 routes, and 11 tranches have more than 10 routes.

[23] The equivalent statistics for the whole sample are very similar.

On average 3.7 bidders submit at least one bid on an auction. The number of bidders ranges between 1 and 8. Fewer bids are submitted on individual routes. On average 2.94 bidders submit a bid for an individual route. The number of bids per route ranges between 1 and 7. A total of 44 bidders submit at least one bid on an auction with three routes or less. Of those, 26 win a contract.

Bidders submit a total of 641 stand-alone bids. The distribution of normalized stand-alone bids resembles a log-normal distribution. The average stand-alone bid equals 13.1 in logarithm which amounts to about 490,000 Pounds. Since bidders are committed by their bids, stand-alone bids define implicitly a combination bid (with value equal to the sum of the stand-alone bids). We call a combination bid "non trivial" when it is strictly less than the sum of the component stand-alone bids. On the auctions with two and three routes a total of 83 non trivial combination bids and 218 trivial combination bids are submitted.[24]

Reasons invoked by the bidders to offer discounts for combinations of routes include the possibility to share spare vehicles and garage overhead costs in general, and more efficient organization and coordination of working schedules. Ignoring trivial combination bids, the discount of a combination bid relative to the sum of stand-alone bids by the same company equals 4.5% on average. The discount amounts to 3.9% with two-route bids, 7.7% with 3-route bids. When all combination bids are included (i.e. those implicitly defined by stand-alone bids), this discount drops to 1.1% on average (1% for two-route bids and 1.6% for 3-route-bids).

The market for bus operators. We denote as a "bidder" any bus operator that is active in the tendered bus services within the local areas of London. As of November 2000, there were 51 independent pre-qualified bidders in the market. After the privatization of the London Buses subsidiaries in 1994, a substantial reorganization and consolidation of the industry took place. Since then, the market has stabilized with a C4 ratio around 70% between late 1996 and 2001.

For each bidder active in the tendered bus services in London, we have a complete history of its garages (openings/first time use for the tendered market and closings, location) since deregulation, as well as its committed fleet for the tendered market on a monthly basis.

[24] A trivial combination bid is generated for each non overlapping combination of routes s and t over which a bidder bid without submitting a bid on route combination $s \cup t$.

The following calculation provides a sense of the censoring present in our data due to the reserve price and the combination bid constraint imposed by LTB. If *all* bidders who ever submitted a bid on a route in a tranche had submitted a bid on all the routes and route combinations in the tranche, we would have 852 stand-alone and 693 combination bids.

Garages are leased on a long term basis or bought, and a typical garage has capacity for 50-100 buses and serves about 8 routes.

A few elements are worth noting. First, asymmetry among bidders is considerable. The average bidder has about four garages in London but the number of garages ranges between 0 and 21 across bidders. For example, in November 2000, a total of 10 bidders had one garage, 4 bidders had two garages, one bidder had 6 garages, one bidder has 7 garages, one bidder had 9 garages, one bidder had 11 garages, one bidder have 13 garages, and one bidder has 21 garages. This size asymmetry is also reflected in the distribution of market shares in our sample.

Second, despite a fairly concentrated market, an active fringe of small bidders seems to be providing a certain level of competition. For our whole sample, "entrants", i.e. bidders without an established garage at the time of the auction, submitted 10.6% of all the bids, and bidders with only one established garage submitted another 15.95% of the bids. In our sample, there was an entrant or a bidder with only one garage bidding on 49.11% of the routes. Such active fringe would make collusion very difficult to sustain.

Money left on the table and bidder participation. A measure of money left on the table is the relative difference between the lowest and second lowest stand-alone bid. Table 1 reveals that the money left on the table equals 13.54%. Thus, stand-alone bids "overpay" by about 110,500 Pounds on average. This suggests that the winning bidder does not know the competitors' bids and is uncertain about the competitors' bid levels.[25]

What determines uncertainty in bids? At the bidder level, costs are determined in part by the actual expenses in capital, labor and fuel incurred in carrying out the contract. But they also depend on the opportunity of using these resources, especially capital, in other ways. There is probably little uncertainty among bidders concerning the expected cost of labor or fuel (there are well functioning markets for these), but opportunity costs may not be known to other bidders. Our interpretation is that *uncertainty in this market is best viewed as stemming from private information about (opportunity) costs.*

An important question for modelling bidding behavior in the London bus routes market is to determine whether cost uncertainty arises at the firm, auction, or route level. In other words, does the opportunity cost vary at the firm level, the garage level or route level? To

[25] As the number of bidders increases, the amount "overpaid" decreases. The money left on the table equals 20.94% when two bids are submitted, 11.51% when three bids are submitted, 9.36% when four bids are submitted and 7.69% when five or more bids are submitted. Even so, with five or more bidders the amount "overpaid" for the average contract equals almost 63,000 Pounds.

examine these questions we decompose the variation in the bid submission decision.

In the next paragraphs, we examine how much of the variation in the decision variable is explained by auction fixed effects, route fixed effects, auction-garage and auction-bidder fixed effects, as well as dead mileage (closest distance from the route to the garage). We focus on bid submission decisions by bidders with an active garage at the time of the auction and on auctions of two and three routes. We are left with 3,358 observations. Due to the large number of explanatory variables, we consider the linear probability model and estimate it using OLS. The empirical model is $y = X\lambda + u$, where $y = 1$ if a bid is submitted and zero otherwise, X denotes a vector of explanatory variables and u denotes the residual.

Table 2 reports our results for several specifications. The individual specifications gradually add more variables to X. A description of these is given in the second column, and their number is given in the third column. The fourth and fifth columns report the R^2 and adjusted $\overline{R}^2$ for the specification. We interpret the increase in the fraction of explained variance as a measure of the importance of the added variables. The last column reports the value of the F- statistic for the test of joint significance of the explanatory variables added relative to the previous model. For example, the test statistic for the hypothesis that auction fixed effects are zero (model (2)) is an F- distributed random variable with (67, 3288) degrees of freedom.

Table 2 Variance Decomposition of the Bid Submission Decision*

Variables Included	#var	R^2	$\overline{R}^2$	F
(1) Dead Mileage, linear and quadratic (DM)	3	0.24	0.24	520.75**
(2) DM+Auction Fixed Effects (AF)	70	0.28	0.27	3.00**
(2') DM+Route Fixed Effects	170	0.30	0.26	1.63**
(3) DM+AF+ Operator Fixed Effects	92	0.31	0.29	4.14**
(4) DM+AF+Depot Fixed Effects	159	0.40	0.37	6.83**

* Auctions with 2 and 3 routes. ** indicates significance at 1% level.

Models (2) and (2') test competing interpretation of the sources of uncertainty common to all bidders: at the auction or at the route level. Model (2') does somewhat worse on the basis of the adjusted $\overline{R}^2$ than model (2) suggesting that there may be little common shocks to bidders at the route level. We tested the null that route fixed effects are zero when auction fixed effects are present. The test statistic is an F- distributed random variable with (100, 3188) degrees of freedom. It is equal to 0.73. We cannot reject the null hypothesis that route fixed effects are zero *once auction fixed effects are accounted for.* We conclude that there are no route level shocks common to all bidders. For this reason, models (3) and

(4) build on model (2).

According to the R^2 in model (4) about 60% of the variation remains unexplained. The unexplained part comes from the remaining uncertainty as to whether a bidder submits a bid on a given route after controlling for dead mileage, garage fixed effects and auction fixed effects. We may interpret this uncertainty as a bidder specific idiosyncracy arising at the route and auction level. Notice also, that the order in which we add variables may affect the contribution to the R^2. We looked at permutations of the order and found no major differences.

The empirical evidence suggests the following origins for the cost uncertainty: First, there is no evidence of cost shocks common to all bidders at the route level after controlling for auction fixed effects since route fixed effects are not significant. Second, a substantial part of the uncertainty in bidders' decisions is explained by bidder asymmetry captured by dead mileage, bidder fixed effects and garage fixed effects. Third, there is considerable residual uncertainty for each bidder arising at the route and auction level.

Summary and conclusions. The evidence presented in this section supports the view that a multi-unit combinatorial first price auction with private values and multi-dimensional private information is a reasonable model for the London bus routes market. We argue these points in turn.

Multi-unit combinatorial: The auction appears the proper level of analysis for this market. First, the temporal simultaneity of the auction for the routes in the same auction, their geographic proximity as well as the existence of combination bids requires that we analyze them at the same time. Second, several elements suggest that inter-auction effects may not be very important. The delay of 10 months between the award date and the start of the contract reduces the role for capacity in this market. In addition, combination bidding is motivated in part by (local) cost synergies among routes, but different auctions tend to cover different geographical areas. The geographic dispersion of the auctions together with local nature of the business reduces the interactions among auctions and bidders. We calculated that an average bidder in our sample bid on an auction every 5 months only.

Private Values: Most of the inputs used by bidders have well-functioning markets. In addition, our bidders are experienced so we expect them to be able to forecast accurately their costs, in the sense that cost forecasts by competitors should not lead to revise their own cost estimates. Finally, the fact that we did not find evidence of common shocks at the route level lends further support to this hypothesis.

Multi-dimensional private information: Our specification is flexible. It allows for cost correlation across routes, but does not assume it. This flexibility seems important in view of the evidence presented in table 2.

5 Estimation Method

This section describes our estimation approach and illustrates it's implementation in practice.

Our estimation approach consists of two steps: In the first step, the bid density is estimated and the implied probability of winning is obtained. In the second step, we use the results of section 3 to infer costs or cost ranges that rationalize the observed bids. Section 5.1 describes the details of our parametric specification of the bid density which takes into account bidder and contract specific covariates. Section 5.2 describes the estimation procedure to infer the parameters of the bid density. Section 5.3 takes the bid density function as given and describes our method to infer costs.

As described in section 3, we observe data on a cross section of auctions $t = 1, \ldots T$. Let $\mathbf{b}^{it}$ denote the bid vector of bidder i submitted for the contracts in auction t. Let $\mathbf{X}^t = (\mathbf{x}^t, \mathbf{w}^{1t}, \ldots, \mathbf{w}^{Nt})$ denote the contract and bidder characteristics of auction t and $\mathbf{X}_s^t$ denote the subset of characteristics relevant for contract s in auction t. Let $\mathbf{w}^{-i,t}$ denote the vector of characteristics for bidders other than bidder i. We sometimes also write $\mathbf{X}^{it} = (\mathbf{x}^t, \mathbf{w}^{it}, \mathbf{w}^{-i,t})$, where superscript i indicates that bidder characteristics are evaluated from bidder i's perspective.

Assumption 1 in section 2 implies that bidder i's bid vector in auction t is stochastically independent from bidder j's bid vector, conditional on observable characteristics. We assume that characteristics $\mathbf{X}^t$ are observable to all bidders and the econometrician. We do not consider bidder or contract heterogeneity that is not observed to the econometrician. Unobserved contract heterogeneity in single unit auctions is studied in Krasnokutskaya (2004). Finally, if there are multiple equilibria, we assume that our data are generated by a single equilibrium selection.

5.1 Bid Density Function

This section describes our specification of the bid density function.

As proposed in Guerre, Perrigne and Vuong (2000), non-parametric techniques can be used to estimate the bid distribution. Non-parametric estimators permit flexibility in the shape of the bid distribution but, at the same time, require many data points to obtain precise estimates. The data requirement increases quickly with the dimensionality of the distribution, as is illustrated in Silvermann (1986). Our bidding model is multi-dimensional and the dimensionality increases further due to the combination bids. Since the size of our data set is small, we consider a parametric specification instead.

Our parametric framework enables us to take covariates into account. In particular, bidder heterogeneity is an important element of London bus auctions. We incorporate bidder heterogeneity into the analysis in three ways: First, we account for efficiency differences between bidders reflected in bidder size and measured by the number of available garages. Second, we take into account the locational specifics of the market by measuring the distance between the garage and the route[26] to account for bidder specific transportation costs. Third, we explicitly distinguish two types of bidders: regular and fringe. Regular bidders have a garage within a eight mile radius of a route in the auction. Fringe bidders do not have a garage yet, or their garage is further than eight miles away from any route in the auction. Distinguishing between regular and fringe bidders allows us to account for behavioral differences between bidders that bid regularly and bidders that bid once on a single contract or on few occasions only.

The statistical model for latent bids $\mathbf{b}^*$ by bidder i on auction t with m^t routes is based on a multi-variate normal density, $\phi(\cdot|\mathbf{X}^{it}, \boldsymbol{\theta})$, with covariates $\mathbf{X}^{it}$ and parameter vector $\boldsymbol{\theta} = (\boldsymbol{\mu}, \boldsymbol{\Sigma})$. We assume that a logarithmic transformation of bids minus the logarithm of the internal cost estimate, **ICE**, is distributed as follows:

$$\ln(\mathbf{b}^*) - \ln(\mathbf{ICE}) \sim \phi(\cdot|\mathbf{X}^{it}, \boldsymbol{\theta})$$

where ϕ denotes a multi-variate normal density and the first element of the parameter vector, $\boldsymbol{\mu}(\mathbf{X}^{it})$ denotes the $(2^{m^t}-1)$ dimensional vector of means of log bids and the second, $\boldsymbol{\Sigma}\left(\mathbf{X}^{it}\right)$, is a $(2^{m^t}-1)$ by $(2^{m^t}-1)$ covariance matrix. We assume that the mean μ_s is a linear function of characteristics,

$$\mu_s(\mathbf{X}^{it}) = \beta_s \cdot \mathbf{X}^{it},$$

where the variables in $\mathbf{X}^{it}$ include bidder i's closest distance from the garage to the start (or ending) point of route (combination) $s \subseteq S^t$ measured in logs, the logarithm of the number of garages of bidder i and an indicator variable that equals one if the bidder is a fringe bidder.

[26] The distance of a bidder to a route is defined as the minimum distance between the route start (or end) point and any garage of the bidder.

Three restrictions on bids characterize the boundaries of the bid support*:* First, latent bids for a combination of items that are greater than the sum of constituent bids are not observed. Second, a stand-alone bid that is larger than a combination bids involving the stand-alone contract is not observed. Third, latent bids above the reserve price supremum are not observed. As in Laffont, Ossard and Vuong (1995), we can incorporate the bid restrictions by defining the observed bid as a function of the latent bid. Specifically, we define the observed bid as equal to the supremum of reserve price, R_s, when the latent bid is not observed. We define the observed bid as equal to the sum of stand-alone bids if this sum is less than the latent bid and if every constituent stand-alone bid in the sum is less than its reserve price. We define the observed constituent bid on a route (combination) as equal to the combination bid involving that route if the latent constituent bid exceeds the combination bid. Formally, we define a mapping $g = (g_1, \ldots, g_s, \ldots, g_S)'$ that transforms latent bids, $\mathbf{b}^* \in \Re_{++}^{2^m-1}$, into observed bids, $\mathbf{b} \in \Re_{++}^{2^m-1}$, that satisfy these constraints, $\mathbf{b} = g(\mathbf{b}^*, \mathbf{R})$. A typical component $s \subseteq S$ of the mapping g is given by:

$$g_s(\mathbf{b}^*, \mathbf{R}) = \min\left(R_s, \min_{s_1, s_2 \subseteq S; s_1 \cup s_2 = s, s_1 \cap s_2 = \varnothing, b^*_{s_1} \leq R_{s_1}, b^*_{s_2} \leq R_{s_2}} (b^*_s, b^*_{s_1} + b^*_{s_2}), \min_{s' \subseteq S} b_{s \cup s'}\right)$$

Henceforth, we use this convention and restrict attention to observed bids. Thus, the support of observed bids by bidder i is given by

$$\mathcal{B} = \left\{ b \in \Re_{++}^{2^m-1} \left| b_s \leq R_s;\ b_{s \cup s'} \leq b_s + b_{s'};\ \text{and } b_s \leq b_{s \cup s'} \text{ for all } s, s' \subseteq S \text{ with } s \neq s' \right. \right\}. \tag{14}$$

and we assume that the supremum of reserve prices are additive, $R_{s \cup s'} = R_s + R_{s'}$ for all $s, s' \subseteq S$ with $s \neq s'$. Notice that all bids in the interior of the support $\mathcal{B}$ are relevant bids as defined in section 3. Irrelevant bids may arise on the boundary of the bid support, for example, when the constituent bid constraint, $b_s \leq b_{s \cup s'}$, is binding.

The total number of bidders at an auction consists of regular and fringe bidders. The set of regular bidders is described by our garage data which include detailed information on all garage locations. The number of fringe bidders is not observed in the data and needs to be estimated. The maximum number of fringe bidders submitting a bid equals three on any auction. We use this number as an estimate for the potential number of fringe bidders. The maximum number of fringe bidders is a superefficient estimator as it converges at a rate faster than our parameter estimates. Thus, the estimate of the number of fringe bidders does not affect the distribution of the remaining parameter estimates asymptotically.

Estimation: There is a large literature on estimation methods of the parameters of a lognormal density function, see Griffiths (1980). Proposed methods include maximum likelihood and the method of moments. Maximum likelihood yields consistent and efficient estimates of the parameters $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ as the number of auctions T gets large. Yet, the likelihood function involves multi-variate integrals. These integrals arise due to the combination, stand-alone and reserve price restriction. Numerical calculation of the integrals and the likelihood function can be computationally intensive. The method of moments provides an alternative estimation method to maximum likelihood that yields consistent estimates as the number of auctions T increases, see Hansen (1982). Numerical calculation of the moment conditions can also be computationally intensive for a multi-variate truncated density. However, simulation estimators (McFadden (1989) and Pakes and Pollard (1989)) provide an elegant solution to this problem.[27] The next section describes our simulated methods of moment estimator for the density of bids.

5.2 A Simulated Method of Moments Estimator

We consider a method of moments estimator that is based on the moments of observed bids. For computational reasons, we consider the logarithmic transformation of observed bids normalized by the internal cost estimate, $\mathbf{B} = \ln(\mathbf{b}) - \ln(\mathbf{ICE})$. Let $(\mathbf{B})^1 = (B_1, \ldots, B_S)$ denote the vector of first moments of log normalized bids, $(\mathbf{B})^2 = ((B_1B_1), (B_1B_2), \ldots, (B_SB_S))$ denote the vector of second moments of log normalized bids, and so on. Notice, that the expected difference between the observed and theoretical moment when evaluated at the true parameter value $\boldsymbol{\theta}_0$ is zero:

$$E\left[(\mathbf{B}^{it})^k - E\left[\mathbf{B}^k | \mathbf{X}^{it}, \boldsymbol{\theta} = \boldsymbol{\theta}_0\right]\right] = 0$$

Given this condition together with some standard regularity conditions, we can adopt the method of moments estimator described in Hansen (1982). Unfortunately, $E\left[\mathbf{B}^k | \mathbf{X}^{it}, \boldsymbol{\theta}\right]$ is the kth moment of a truncated multi-variate normal random variable, which is numerically time-consuming to calculate. We solve the integration problem by replacing the difficult to calculate expected value with a simulated, unbiased estimate. To see this, notice that the expected kth order moment of the observed bid can be written as

$$E\left[\mathbf{B}^k | \mathbf{X}^{it}, \boldsymbol{\theta}\right] = \int \cdots \int \left(\log(\overline{g}(\boldsymbol{\varepsilon}, \mathbf{R}^t))^k \frac{\phi(\boldsymbol{\varepsilon} | \mathbf{X}^{it}, \boldsymbol{\theta})}{\varphi(\boldsymbol{\varepsilon})} \right) \varphi(\boldsymbol{\varepsilon}) d\varepsilon_1 \cdots d\varepsilon_{S^t}$$

by multiplying and dividing the integrand by the multi-variate importance function $\varphi(\cdot)$, which we assume equal to the product of univariate normal densities, $\varphi(\boldsymbol{\varepsilon}) = \Pi_{s \subseteq S^t} \psi(\varepsilon_s)$ and

[27] Methods based on simulating the likelihood function are less attractive because the simulation error enters non-linearly and may bias the estimates.

where $\psi(\cdot)$ denotes a univariate normal density with unit variance. The mapping $\overline{g}(\boldsymbol{\varepsilon}, \mathbf{R}^t)$ normalizes bids by using the internal cost estimate and takes the reserve price, stand-alone and combination bid restriction into account. So, a typical component of the mapping $\overline{g}$ is given by $\overline{g}_s(\boldsymbol{\varepsilon}, \mathbf{R}^t) = g_s(\exp(\boldsymbol{\varepsilon}) \cdot \mathbf{ICE^t}, \mathbf{R}^t)/ICE_s^t$.

Given a fixed set of L random draws, $\widehat{\boldsymbol{\varepsilon}} = \left(\widehat{\boldsymbol{\varepsilon}}^1, \ldots, \widehat{\boldsymbol{\varepsilon}}^L\right)$, from the multi-variate importance function $\varphi(\cdot)$ for each bidder and auction, we define the estimate, $\widehat{B^k}$, for bidder i on auction t as:

$$\widehat{B^k}(\mathbf{X}^{it}, \boldsymbol{\theta}, \widehat{\boldsymbol{\varepsilon}}^{it}) = \frac{1}{L} \sum_{l=1}^{L} \left[\log(\overline{g}\left(\widehat{\boldsymbol{\varepsilon}}^{itl}, \mathbf{R}^t\right))^k \frac{\phi(\widehat{\boldsymbol{\varepsilon}}^{itl} | \mathbf{X}^{it}, \boldsymbol{\theta})}{\varphi(\widehat{\boldsymbol{\varepsilon}}^{itl})} \right]$$

Observe that the estimate $\widehat{B^k}(\mathbf{X}^{it}, \boldsymbol{\theta}, \widehat{\boldsymbol{\varepsilon}}^{it})$ is continuously differentiable in the parameter vector θ and it is an unbiased estimator of $E\left[\mathbf{B}^k | \mathbf{X}^{it}, \boldsymbol{\theta}\right]$,

$$\widehat{B^k}(\mathbf{X}^{it}, \boldsymbol{\theta}, \widehat{\boldsymbol{\varepsilon}}^{it}) = E\left[\mathbf{B}^k | \mathbf{X}^{it}, \boldsymbol{\theta}\right] + \widehat{v}^k.$$

Under regularity conditions which are satisfied here, McFadden (1989) and Pakes and Pollard (1989) show that a method of moments technique is still appropriate to estimate $\boldsymbol{\theta}$ when the estimate $\widehat{B^k}$ replaces the theoretical analogue. We consider two sets of moment conditions. First, the first moment of observed bids:

$$\mathbf{M}_1(\boldsymbol{\theta}) = \sum_{i,t} \sum_{s \subseteq S^t} \left[B_s^{it} - \widehat{B_s^1}(\mathbf{X}^{it}, \boldsymbol{\theta}, \widehat{\boldsymbol{\varepsilon}}^{it}) \right] \cdot \mathbf{W}_s^{it1}$$

where $\mathbf{W}_s^{it1}$ is a row vector of instruments for the first moment of the normalized bid B_s^{it}. The second set of moment conditions is based on the second moment of observed bids and accounts for the correlation between (combination) bids:

$$\mathbf{M}_2(\boldsymbol{\theta}) = \sum_{i,t} \sum_{s,s' \subseteq S^t} \left[(B_s^{it} \cdot B_{s'}^{it}) - \widehat{B_{ss'}^2}(\mathbf{X}^{it}, \boldsymbol{\theta}, \widehat{\boldsymbol{\varepsilon}}^{it}) \right] \cdot \mathbf{W}_{ss'}^{it2}$$

where $\mathbf{W}_{ss'}^{it2}$ is a row vector of instruments for the second moment of the normalized bid $B_s^{it} \cdot B_{s'}^{it}$. The method of moment estimation technique is based on the restriction that the moment condition is uncorrelated with the exogenous data $\mathbf{W} = (\mathbf{W}^1, \mathbf{W}^2)$. Let $\mathbf{M}$ denote the column vector of moment conditions,

$$\mathbf{M}(\boldsymbol{\theta}) = (\mathbf{M}_1(\boldsymbol{\theta}), \mathbf{M}_2(\boldsymbol{\theta}))',$$

where the dimensionality of the vector of moment conditions is at least as large as there are parameters in $\boldsymbol{\theta}$. An estimate $\widehat{\boldsymbol{\theta}}$ is chosen to minimize a quadratic distance measure

$$\mathbf{M}(\boldsymbol{\theta})'\mathbf{A}\mathbf{M}(\boldsymbol{\theta})$$

for some positive definite matrix $\mathbf{A}$. A preliminary estimate $\widehat{\boldsymbol{\theta}}_0$ is obtained by setting $\mathbf{A}$ equal to the identity matrix. Then, a new weighting matrix is calculated as the inverse of the sample variance of the individual moment conditions, $\widehat{\mathbf{A}} = \left[\frac{1}{\sum_t N^t} Var\left[\mathbf{M}(\widehat{\boldsymbol{\theta}}_0)\right]\right]^{-1}$, where N^t equals the number of bidders in auction t. A second and final estimate, $\widehat{\boldsymbol{\theta}}$, is then obtained from the use of this moment condition. The estimate $\widehat{\theta}$ is consistent and $\sqrt{\sum_t N^t}\left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right)$ is asymptotically normally distributed with zero mean and covariance matrix

$$\left(1+\frac{1}{L}\right)\left(E\frac{\partial}{\partial\boldsymbol{\theta}}\mathbf{M}'\widehat{\mathbf{A}}E\frac{\partial}{\partial\boldsymbol{\theta}}\mathbf{M}\right)^{-1},$$

see for example Pakes and Pollard (1989).

A Monte Carlo study revealed that the estimator is well behaved even for small number of observations. Moreover, the first two moments, $k = 1, 2$, are sufficient to identify the parameter vector $\boldsymbol{\theta}$.

5.3 Inference of Costs

This section describes our technique to infer costs for a typical auction t. For simplicity of exposition we omit the auction superscript. The primitive for our inference procedure is the density of observed bids. The bid density allows us to obtain an expression for the density of opponents' low bids and the probability of winning. Costs and the cost distribution function can then be inferred based on the first order condition for optimal bids.

Let $h(b_1^i, ..., b_s^i, ..., b_S^i|\widehat{\theta}, \mathbf{X}^i)$ denote the estimated probability density function evaluated at bids by bidder i on all the subsets of S conditional on $\mathbf{X}^i$. For simplicity of exposition, we denote the secret reserve prices of the auctioneer as bids submitted by bidder zero $\left(b_s^0\right)_{s\subseteq S}$, and write the density of reserve prices as $h(b_1^0, ..., b_s^0, ..., b_S^0|\widehat{\theta}, \mathbf{X}^0)$. This notational convention allows us to include the reserve prices in the set of bids by i's opponents. The probability that bidder i's bid (vector) $\mathbf{b}^i$ wins exactly contract s conditional on $\mathbf{X}^i$, $G_s(\mathbf{b}^i|\widehat{\theta}, \mathbf{X}^i)$ can be written as a function of the density of the cheapest allocation of the routes in S among bidder i's opponents conditional on $\mathbf{X}^i$, which we denote as $h_{(1)}(.|\widehat{\theta}, \mathbf{X}^i)$. Notice further that the density $h_{(1)}(.|\widehat{\theta}, \mathbf{X}^i)$ can be expressed directly as a function of the bid densities $h(.|\widehat{\theta}, \mathbf{X}^j)$ for $j = 0, \ldots, N$. The analytical expressions involves multi-dimensional integrals which are complex to calculate numerically.

Following Judd (1998), we solve the described integration problem using Monte Carlo integration methods. The method is based on the law of large numbers and can be explained as follows: For each bidder $j \neq i$ we draw a bid (vector) from the density $h(.|\widehat{\theta}, \mathbf{X}^j)$ conditional on the characteristics $\mathbf{X}^j$. Additionally, a random reserve price (vector) is drawn from the reserve price distribution and included in the set of bids. We then determine the cheapest allocation of the contracts in S among bidder i's opponents $(\underline{\mathbf{b}}^{-i}|\widehat{\theta}, \mathbf{X}^i) = (\underline{b}_1^{-i}, ..., \underline{b}_S^{-i}|\widehat{\theta}, \mathbf{X}^i)$ as defined in equation (1). We repeat this exercise L times by repeatedly drawing bids and determining the cheapest allocation of the routes in S among bidder i's opponents. The "pseudo data" of cheapest allocations, $(\underline{\mathbf{b}}^{-il}|\widehat{\theta}, \mathbf{X}^i)_{l=1}^L$, is then used directly to approximate the probability that bidder i wins exactly route s with the bid $\mathbf{b}^i$ contained in the interior of the support $\mathcal{B}$,[28] $G_s(\mathbf{b}^i|\widehat{\theta}, \mathbf{X}^i)$. The empirical frequency of this event is given by:

$$G_s(\mathbf{b}^i|\widehat{\theta}, \mathbf{X}^i) = \frac{\sum_l 1_{\{\text{bid } \mathbf{b}^i \text{ wins exactly route } s| \ (\underline{\mathbf{b}}^{-il}|\widehat{\theta}, \mathbf{X}^i)\}}}{L},$$

where $1_{\{x\}} = 1$ if x is true and 0 otherwise. By the law of large numbers, the approximation error vanishes as L increases.[29] The partial derivative $\partial G_s / \partial b_t$, can be calculated numerically by using one sided differences with $\varepsilon(L)$ appropriately chosen.[30] The numerical difference yields,

$$\frac{\partial G_s(\mathbf{b}^i|\widehat{\theta}, \mathbf{X}^i)}{\partial b_t} = \frac{G_s(b_1^i, ...b_t^i, ..., b_S^i|\widehat{\theta}, \mathbf{X}^i) - G_s(b_1^i, ...b_t^i - \varepsilon, ..., b_S^i|\widehat{\theta}, \mathbf{X}^i)}{\varepsilon}.$$

Section 3 describes how to obtain the cost range associated with the bid vector $\mathbf{b}^i$ based on the Kuhn-Tucker conditions for optimal bids. The Kuhn Tucker conditions are fully characterized with the estimates of G_s, $\partial G_s / \partial b_t$ in hand. As is explained in section 3, the cost range can be either a singleton, or a path or a higher dimensional area. Moreover, the cost range can be calculated by varying the Lagrange multiplier(s) between 0 and ∞. Finally, we observe that our numerical approximation entails an error and the error becomes negligible as L increases.

[28] For a bid $\mathbf{b}^i$ on the boundary of the support, the probability of winning equals the limiting winning probability of a sequence of interior bids that converges to $\mathbf{b}^i$. By continuity of the assumed bid density and the randomness of the reserve price, the winning probabilities in the sequence will converge to the winning probability at the boundary. Thus, we may approximate the winning probability of a boundary bid with the winning probability of a close by interior bid.

[29] A smooth approximation of G_s is obtained by employing a kernel estimator, or similar smoother, on the pseudo data, see Silverman (1986).

[30] The derivative on the boundary of the bid support can be approximated by the derivative evaluated at close by interior bids.

6 Estimation Results

This section describes our estimates. We report the parameter estimates of the bid density functions in section 6.1. Section 6.2 illustrates the bidder objective function and optimal bids. Finally, section 6.3 illustrates our estimates for cost synergy bounds and assesses bounds on potential efficiency losses due to the auction rule.

6.1 Estimates

Our data do not include information on the auctioneer's reserve prices. Reserve prices are kept secret and bidders do not know the reserve prices at the time of bid submission.[31] To account for the reserve price in the estimation, we presume that a secret reserve price is drawn independently for each route from a uniform distribution with interval support. The boundary points of the reserve price support are assumed linear in the internal cost estimate. Reasonable candidates for the linear coefficient can be inferred from the ratio of accepted bid to the internal cost estimate. Excluding two single route auctions which are outliers[32] an upper bound to this ratio equals 1.45. We use this number as an upper bound for our analysis. As a lower bound for the secret reserve prices we use the internal cost estimate. Altering the reserve price rule would not affect our statistical model of bids. In fact, in the estimation of the bid density, we lower the upper reserve price bound from 1.45 to 1.3 times the internal cost estimate to diminish possible errors stemming from the presumed reserve price rule.[33] For our cost estimates, the effect of the presumed secret reserve price is to smooth the simulated winning probability for bids that approach the upper reserve price bound, and thus to reduce the estimated bidders' markup over costs. The reason for the reduced markup is that the secret reserve price has a similar effect on the calculus of bidding as adding an additional bidder. It enhances competition in the auctions and results in reduced markups over costs.

Parameter specification and instruments: There are a number of natural restrictions to impose on the way the parameters enter the bid density of regular bidders. The parameters should be invariant with respect to permutations of the indices of sets of routes. For

[31] Our data include a total of five rejected bids.

[32] The two excluded auctions have accepted winning bids that are 52 and 108 percent above the internal cost estimate. These auctions consist of single contracts with an internal cost estimate of less than 50,000 Pounds. The internal cost estimate is two standard deviations below the mean internal cost estimate and possibly erroneously measured.

[33] A total of six bid observations become truncated due to the reduced upper boundary point.

the mean the invariance implies that dead mileage and other covariates should have the same effect for each contract, or formally $\beta_s = \beta_\tau$, for all s, τ such that $|s| = |\tau|$. For the covariance matrix the restriction implies that the matrix $\mathbf{\Sigma}$ should be symmetric and that the covariance between routes one and two should equal the covariance between routes one and three, and so on. Formally, the restriction can be stated as $\alpha_{s,s} = \alpha_{\tau,\tau}$, $\alpha_{s,\tau} = \alpha_{\tau,s}$, $\alpha_{s,u} = \alpha_{\tau,u}$ and $\alpha_{u,s} = \alpha_{u,\tau}$ for all s, τ such that $|s| = |\tau|$, and $s, \tau \neq u$.

In the estimation of the mean, we want to control for contract and bidder heterogeneity. Contract heterogeneity is accounted for by dividing bids by the internal cost estimate. As bidder specific explanatory variables for regular bidders we include linear and quadratic expressions of the logarithm of dead mileage and the logarithm of the number of garages of bidder i. We include a FRINGE dummy variable that equals one if the bid is submitted by a fringe bidder and zero otherwise. Typically, fringe bidders do not have a garage established at the time of bid submission, and we adopt the convention that the logarithm of the dead mileage variable and the number of garage variable is equal to zero for fringe bidders. Thus, the coefficient on the FRINGE dummy takes into account the mean of fringe bids and is measured relative to a regular bidder with no garages and dead-mileage equal to zero. This yields the following specification for μ_s^i:

$$
\begin{aligned}
\mu_s^i = {} & \beta_0 + \beta_1 \text{DEAD-MILEAGE-}i\text{-}s + \beta_2 \left(\text{DEAD-MILEAGE-}i\text{-}s\right)^2 \\
& + \beta_3 \text{NO-GARAGES-OF-}i + \beta_4 \text{FRINGE}
\end{aligned}
$$

For a single route auction the covariance matrix $\mathbf{\Sigma}$ equals the parameter λ_1. For multi-route auctions we specify the elements $\alpha_{s,\tau}$ of the covariance matrix $\mathbf{\Sigma}$ as follows:

$$
\begin{aligned}
\alpha_{s,\tau} = {} & \lambda_1 1_{\{s=\tau, \text{ and } s \text{ is a single route}\}} + \lambda_2 1_{\{s \neq \tau \text{ and } s \cap \tau = \emptyset\}} + \lambda_3 1_{\{s \neq \tau \text{ and } s \cap \tau \neq \emptyset\}} \\
& + \lambda_4 1_{\{s=\tau, \text{ and } s \text{ is a combination route}\}}
\end{aligned}
$$

The first and last constant account for diagonal elements in $\mathbf{\Sigma}$, while the second and third constant account for off diagonal elements. The first constant accounts for the variance of a single route bid and the last constant accounts for the variance of a combination bid. We distinguish two off-diagonal effects depending on whether contracts s and τ have a non-empty intersection or not.

As instruments for the moment condition any of the exogenous data are admissible. These include all bidder and auction specific variables on each auction and the powers of these variables. The total number of instruments has to equal at least the number of parameters. For the mean of route combination s on auction t for bidder i we select the following five instruments: a constant, the dead mileage of regular bidder i to route (combination) s, the dead mileage of bidder i to route (combination) s squared, the number of garages of bidder i and a dummy variable that equals one if the bid is submitted by a fringe bidder. For the second moment between bids b_s and b_τ for bidder i we select four dummy variables as instruments. The dummies are the indicator functions defined in the specification of $\alpha_{s,\tau}$.[34] The total number of instruments equals the number of parameters in the model, which guarantees identification.

As described in section 4, we estimate the parameters using the method of simulated moments.

The estimates are reported in Table 3. The covariance estimates reveal the following: The covariance estimate for bids on distinct routes, or "unrelated" routes, is positive and equals 0.09. It is significantly different from zero. The covariance estimate for "related" routes, that is for a combination of routes and a bid on a constituent stand-alone route, is negative and equals -0.04. It is significantly different from zero.

An alternative measure of bid correlation is the correlation coefficient. Our parameter estimates imply a positive correlation between two stand-alone bids with a correlation coefficient of about 0.21. A bid for a combination of routes and a constituent stand-alone bid are negatively correlated with a correlation coefficient of about -0.17. Bids for "related" routes may exhibit a different correlation than bids for "unrelated" routes if there is a possible lack of synergies, or if there are strategic effects present. We will explore these alternative explanations below. The variance estimate of a log stand-alone bid equals 0.42. The variance estimate of a combination bid is smaller and equals 0.11. The difference in the variance estimate for stand-alone and combination bids is significant.

The estimates reveal the following for the mean regular bid: The bidder specific variables have the expected effects: Firms with a garage located further away from the route submit higher bids than firms with a garage close by. The distance variable is significant. At the sample average values of explanatory variables, a one percent increase in distance increases the bid by 0.68 percent. Larger firms bid less aggressively. A one percent increase in the

[34]Specifically, the first dummy equals one when $s = \tau$, and s is a single route, and zero otherwise. The second dummy equals one when $s \neq \tau$ and $s \cap \tau = \emptyset$, and zero otherwise. The third dummy equals one when $s \neq \tau$ and $s \cap \tau \neq \emptyset$, and zero otherwise. And the fourth dummy equals one when $s = \tau$, and s is a combination route, and zero otherwise.

Table 3 Parameter Estimates for the Distribution of Bids*

Observations	1525	1525	1525
Constant	-0.2818	-0.2799	0.5616
	(0.449)	(0.5970)	(0.586)
Dead-Mileage-i	0.2265	0.2259	0.2362
	(0.032)	(0.059)	(0.068)
Dead-Mileage-i-SQ	-0.1428	-0.1413	-0.5611
	(0.250)	(0.374)	(0.369)
No-Garages-i	0.2236	0.2234	-0.0256
	(0.005)	(0.008)	(0.126)
Fringe	1.1791	1.1771	0.2494
	(0.447)	(0.594)	(0.624)
λ_1	0.4168	0.4185	1.5732
	(0.101)	(0.1223)	(3.737)
λ_2	0.0865	0.0831	0.6284
	(0.041)	(0.038)	(1.4649)
λ_3	-0.0352	-0.0337	-0.0577
	(0.017)	(0.014)	(0.221)
λ_4	0.1063	0.1019	0.5277
	(0.053)	(0.041)	(1.277)
Available Capacity-i		0.0010	
		(0.060)	
λ_5			-0.1191
			(0.168)

* All continuous explanatory variables are in logarithm. Standard errors are in parenthesis.

number of garages increases the bid by 0.23 percent. Fringe bidders submit higher bids than non-fringe bidders. The difference in log bids between a fringe and a non-fringe bidder equals 1.18.

Synergies across auctions: The covariance estimates indicate that (stand-alone) bids submitted from the same garage are positively correlated. It may be possible that the bid correlation extends beyond the current auction as bids from the same garage submitted at different auctions may be related as well. For example, it may be the case that a bidder with garage capacity already committed to earlier won contracts may be reluctant (or eager) to win another contract. As is shown in Jofre-Bonet and Pesendorfer (2003) the inter auction effects can be measured consistently by using a dynamic bidding game. In London bus auctions considerable time elapses between auctions and consecutive auctions typically consist of routes located in distinct geographic areas of London. So, we may expect that the intertemporal effects are less important. However, we may examine whether inter auction effects are present.

The second column in Table 3 reports estimates in which the capacity committed to routes won in prior auctions is included as an additional explanatory variable. The capacity variable is constructed as the total number of buses already committed at the garage divided by the maximum number of buses ever in operation at this garage. As is evident in Table 3, the coefficient of the capacity variable is not significantly different from zero and the inclusion of the additional variable does not affect the other variables' coefficients significantly. The evidence suggests that intertemporal capacity effects may not be important for London bus route auctions, or that the capacity at the garage is a poor measure of intertemporal effects. It is also possible that the auctioneer when deciding which routes to offer in an auction takes potential synergies already into account and offers the routes with potential synergies in one auction. Indeed, the auctioneer tends to offer routes within a geographic area at the same time which may explain the absence of intertemporal effects.

Variance in bids: Our empirical specification normalizes bids by the internal cost estimate which yields a standard deviation of log bids that varies proportionally with the internal cost estimate. It may be possible that the bid standard deviation is of a richer functional form that goes beyond this proportional formula. For example, it may be the case that the standard deviation of log bids is larger on small than on large routes. To account for this possibility, we consider an augmented specification in which the variance of log bids (normalized by the internal cost estimate) is multiplied with the internal cost estimate to the power of λ_5. If the coefficient λ_5 is different from zero, then the standard deviation of normalized bids can shrink or increase with the internal cost estimate. If the coefficient λ_5 is not different from zero, then the richer specification coincides with the specification in column one.

The estimates of the augmented variance model are reported in column three of Table 3. The coefficient for the internal cost estimate is negative but not significantly different from zero. We cannot reject the null that the coefficient λ_5 is equal to zero. Thus the augmented variance specification is rejected vis-a-vis the specification in column one. The rejection of the augmented variance model may indicate also that our data are not sufficiently rich to characterize the dependence of the variance on covariates and auction characteristics in more detail.

Goodness of fit: An economically relevant measure to the auctioneer is the procurement cost of the final allocation. We use this measure to assess the goodness of fit of our estimates. The procurement cost is defined as the cheapest bidder bid allocation across all routes in the auction. We look at bids only, and leave the auctioneer's secret reserve prices aside in the calculation of the procurement cost. To control for auction heterogeneity, we divide the procurement cost realization by the internal cost estimate. We define our goodness of fit measure as the sample average ratio of the procurement cost realization to the internal cost estimate.

In the data, the procurement cost average equals 88.28 percent of the internal cost estimate with a standard deviation of 22.33 percent. In order to calculate the model predicted procurement cost, we take 10,000 simulation draws and determine the average predicted cost across all auctions for each simulation draw. The simulated measure is lower than the observed measure and equals 84.56. The standard deviation of the simulated measure equals 2.36 percent. We cannot reject the null that the two means are equal at the 90% confidence interval. We can conclude that our estimates predict the procurement cost reasonably well.

We also examined whether differences between the simulated and observed procurement cost depends on the internal cost estimate. Dividing our data into small and large auctions, grouped by using the criterion whether the internal cost estimates is below or above the median value, reveals that our estimates slightly over-predict the procurement cost for small contracts and under-predicts it for large contracts. Yet, the relationship is weak. A regression revealed that the internal cost estimate accounts for 2 percent in the variation of the ratio of observed to simulated procurement cost only.

Bids on the boundary of the bid support indicate that one or more bidding constraints are binding. As is explained in section 3, a binding bidding constraint does not allow us to fully identify the cost vector. Instead, the cost vector is partially identified only.

Partial identification is an important element of our application. We illustrate this issue for one and two route auctions: For single route auctions, we find that no constraint is binding on 98 of 551 observations, or 18% of all observations. The remaining bid observations exceed the reserve price supremum, and partial identification emerges thus on 82% of all single route auction observations.

For two route auctions, no constraint is binding on 4 out of 210 observations. Two percent of all observations yield a fully identified cost vector. Partial identification arises due to distinct bidding constraints. Exactly one bid exceeds the reserve price supremum on 7% of all observations, exactly two bids exceed the reserve price supremum on 10% of all observations, and the combination bid constraint is binding on 14% of all observations.

For the remaining observations all reserve price constraints are binding and a lower bound to the cost vector is identified only.

The next section illustrates the bidder objective function and the optimal bid choices.

6.2 Illustration of the Estimates

With the bid density estimates at hand a bidder's decision problem can be illustrated. We select a hypothetical two route auction containing the route with the median internal cost estimate and assume that both routes have an internal cost estimate equal to the sample median internal cost estimate. The number of potential bidders and all other covariates are taken from an auction observation of a typical bidder with 17 potential opponents. The cost for the stand-alone routes are assumed identical and equal to 94% of the internal cost estimate. The cost for the combination of the two routes is assumed to equal twice the cost of a stand-alone route. We take a random sample of bids from the estimated bid density and the secret reserve price distribution. Then, we determine the bidder´s chances of winning. By re-sampling $5,000$ times we obtain a discrete probability distribution function of winning probabilities as a function of the own bids. To facilitate the graphical illustration, we smooth the chances of winning using a kernel estimator.

[Figure 1 about here]

Figure 1 depicts the bidder's objective function on the vertical axis varying the combination and stand-alone bids on the horizontal axis. The combination bid on the horizontal axis is rescaled by the factor of 1/2 to make the combination bid levels comparable with the stand-alone bid. The solid line depicts the value of the objective function varying the combination bid and holding the stand-alone bids fixed at the optimal levels. The dashed line illustrates the value of the objective function varying both stand-alone bids simultaneously, i.e. $b_1 = b_2$, and holding the combination bid fixed at the optimal level. The solid line illustrates that the value of the objective function increases initially, and then decreases, as the combination bid increases. At a combination bid of about 3 million Pounds, the combination bid constraint becomes binding, and the objective function is flat from then onwards. The peak in the objective function occurs to the left of the flat area illustrating that the optimal combination bid is less than the sum of stand-alone bids. The dashed line increases for most of the range. At a stand-alone bid of about 1.3 million Pounds, the combination bid constraint stops binding, and the objective function has a kink increasing more steeply from then on. Eventually, the objective function becomes flat as the reserve

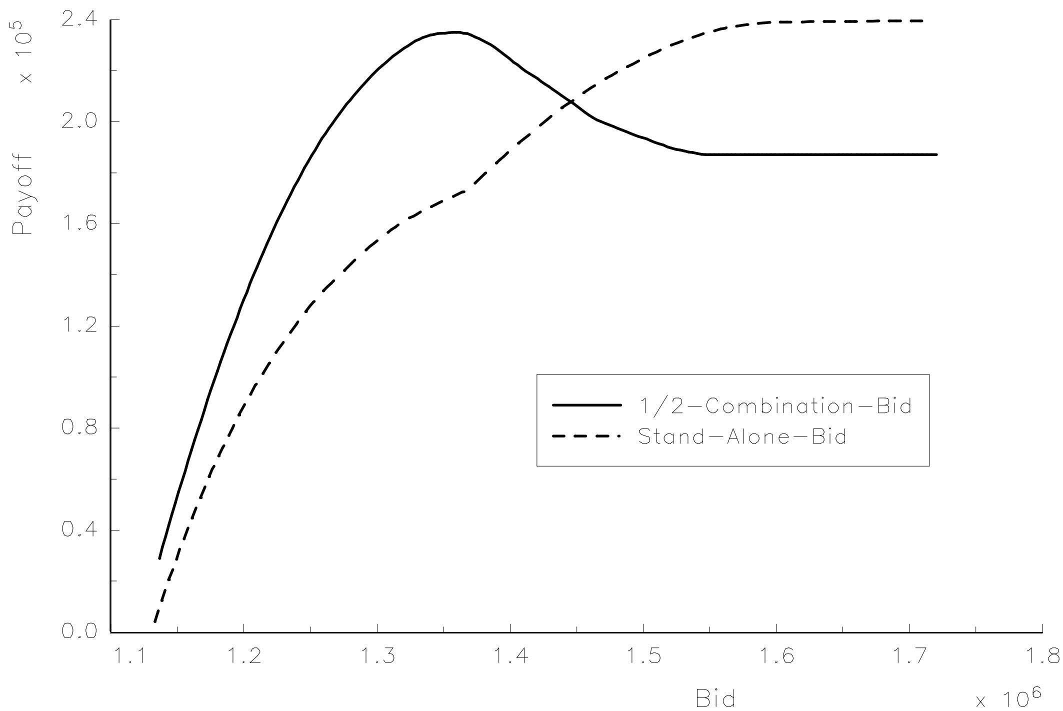

Figure 1: Objective Function

price constraint becomes binding. The optimal stand-alone bids arises at about 1.5 million Pounds which occurs shortly before the flat area at which the reserve price constraint is binding.

Figure 1 illustrates also that the objective function appears to be well behaved with a single maximum. Our identification and inference arguments require that bidders' bids correspond to the global maximum of their expected payoff functions and are thus optimal bid choices given the opponents' bids. As the first order condition is necessary but not sufficient for equilibrium, we need to verify that bids correspond to a global maximum of the estimated payoff function before proceeding with the inference based on the first order condition.[35] As a check that the global maximum requirement is satisfied for our data and estimated bid distributions, we conduct the following exercise. For all observed bid vectors we determine whether these bids correspond to a global maximum of the estimated expected payoff function. To do so, we numerically calculate the optimal bid choice for the inferred cost estimate for every observation in single and two route auctions. An exact match between the optimal bid choice and the observed bid choice cannot be expected due to the numerical approximation error in the cost inference and in the optimal bid calculation. Yet, for small approximation errors, we may expect that the optimal bid vector is close to the observed bid vector.[36] We examine the issue for single route and two auctions. We find that for 99 percent of our observations, every component of the numerically calculated optimal bid vector lies within a one percent radius of the observed bid.[37] This numerical validation is no proof but it gives us sufficient confidence to proceed to estimate cost synergies.

Illustration of the optimal bid choices varying costs: Earlier we argued that even in the absence of cost savings a combination bid may be lower than the sum of stand-alone bids when there is not too much correlation in the value of the cheapest allocation among opponents across contracts. We may illustrate the magnitude of the estimated correlation in opponents' low bids by taking random samples repeatedly from the bid density

[35] For single unit auctions, Guerre, Perrigne and Vuong (2000) show that the first order condition together with the condition that the bid density has a monotone hazard rate provide necessary and sufficient condition for optimality of each bidder's best response. For multi-unit auctions, we are not aware of a condition on the bid density that would guarantee the sufficiency of the first order condition. Instead, we compare the observed bid vectors to the bid vectors that maximize the estimated expected payoff function.

[36] The empirical objective function can be viewed as the true objective function plus a computational error. The computational error comes from the numerical calculation of the inverse matrix of an already numerically approximated gradient matrix. When the computational error is small, the solution will be close to the observed bid by the theorem of the maximum.

[37] The one percent threshold is not reached for one single-route and for eight two-route observations.

for the hypothetical auction and determining the low bid for each sample. We find that the correlation between low stand-alone bids is positive and equals 0.06. The correlation between the low stand-alone bid and the low combination bid is positive and of larger magnitude. It equals 0.36. The positive correlation involving the low stand-alone bid and the low combination bid is in part induced by the combination bid constraint which requires that a combination bid is at most as large as the sum of stand-alone constituent bids. For other routes characteristics and auctions we found correlations of similar magnitudes. The small amount of correlation for stand-alone bids suggests that even without cost savings we may expect that a bidder's combination bid may be less than the sum of stand-alone bids. Next, we illustrate the profit maximizing bids for a range of cost draws.

[Figure 2 about here]

Figure 2 depicts optimal bid choices varying the bidder's cost of the route contracts. We consider here the same auction observation and bidder as in Figure 1. We make the following assumption about costs: The cost for the stand-alone routes are identical and the cost for the combination of the two routes equals twice the cost of a stand-alone route. The solid line denotes the combination bid divided by two, the dashed line equals the stand-alone bid, and the dotted line is the 45 degree line. The profit maximizing bids are calculated at a set of cost points and a line is drawn to connect the solutions. Due to numerical calculation errors arising in the construction of the winning probabilities and also in the optimization routine, the resulting line is not smooth.

Figure 2 shows that the combination bid is less than the sum of stand-alone bids for all the cost range. The difference between the sum of stand-alone bids and the combination bid decrease gradually from 380,000 to 30,000. The relative difference, normalized by the cost, ranges between 30 and 45 percent. Figure 2 also illustrates the bidder's markup as the difference between the bid and the 45 degree line. The markup decreases monotonically as the cost increases. The markup of the combination bid ranges between 2% and 50% of the cost.

As described earlier, on average, the relative difference between the sum of stand-alone bids and the combination bid in the data equals 1.1%, which is substantially smaller than the hypothetical difference in the absence of cost synergies depicted in Figure 2. The low observed difference suggests that there could be lack of cost synergies for London bus routes. We will assess this hypothesis in more detail next.

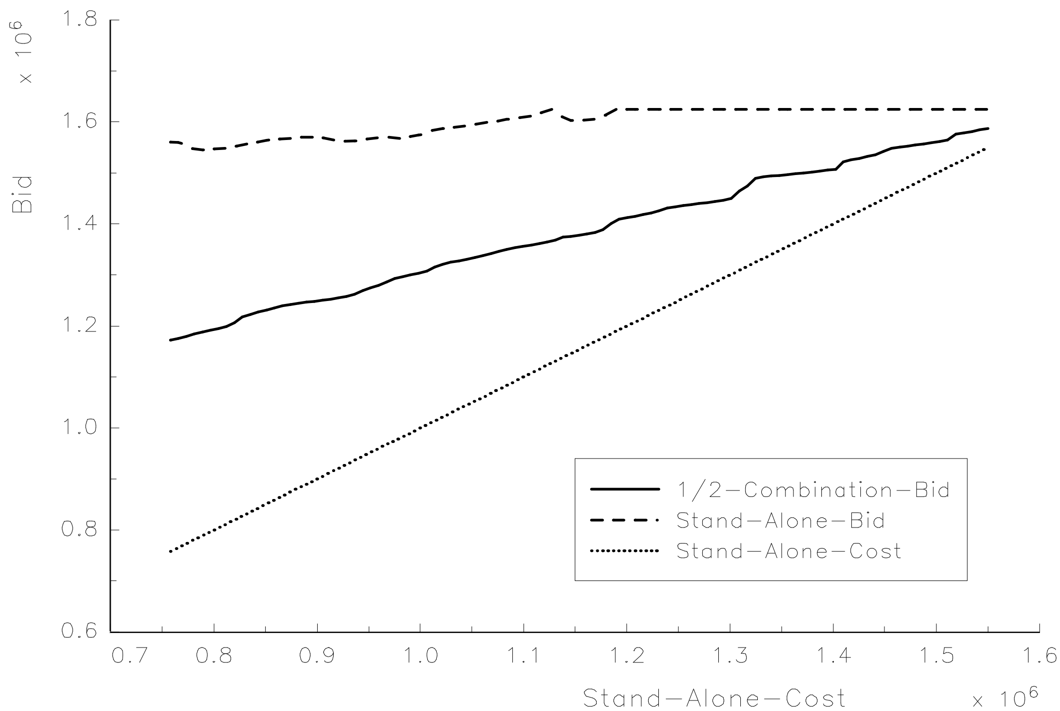

Figure 2: Optimal Bids

6.3 Synergy Estimates

The cost synergy can be inferred based on the inverse of the estimated bid mapping as is explained in section 3. Let σ^{it} denote the cost synergy measure for bidder i on auction t. It equals the difference between the sum of stand-alone costs minus the cost for the combination of routes and is normalized by the absolute value of the sum of stand-alone costs:

$$\sigma^{it}_{s\tau} = \frac{c^{it}_s + c^{it}_\tau - c^{it}_{s\cup\tau}}{|c^{it}_s + c^{it}_\tau|} \quad \text{for } s, \tau \subseteq S^t$$

We take the absolute value of the sum of costs in the normalization as the sum of cost estimates may be negative.[38] The synergy measure is un-determined when a single stand-alone bid is submitted only as there are two degrees of underidentification reflected in $c^{it}_{s\cup\tau}$ and c^{it}_τ. For this reason, we omit observations with isolated stand-alone bids and calculate the synergy measure only for those observations for which at least two stand-alone bids are submitted. As is explained in section 3, point identification of costs is achievable if a (full set of) non-trivial combination bid(s) is submitted in addition to stand-alone bids. When a trivial combination bid is submitted, we evaluate the synergy measure by setting the Lagrange multiplier of the combination bid constraint to zero. The resulting synergy measure provides an upper bound on the magnitude of the cost synergy $\sigma^{it}_{s\tau}$.

For two route auctions there are a total of 57 bids with at least two stand-alone bids. We omit 67 observations in which one stand-alone bid is submitted only. A total of 25 of the 57 observations have a combination bid less than the sum of stand-alone bids. The inferred cost synergy measure is positive on 24 of 57 occasions. The median cost synergy measure is negative and equals -0.11. The estimates imply that the route combination costs at least 11% more than the individual routes separately.

A similar picture emerges when we consider three route auctions. Again, we include bid observations to calculate the synergy measure $\sigma_{s\tau}$ when the bids satisfy one of these three conditions: (i) the submitted bids do not violate any constraint; (ii) the combination bid constraint between routes s and τ is binding but no other constraint is binding; (iii) the stand-alone bids b_s and b_τ are submitted and no stand-alone or combination bid involving the third route is submitted (all other bids are above the reserve price bound). For three route auctions there are a total of 46 observations satisfying the criterion. A total of 36 of the 46 observations have a combination bid less than the sum of stand-alone bids. To

[38] Negative estimates of the opportunity costs for bus operators could arise due to costs of stopping the bus operation completely and laying off the workforce, or due to long term contractual costs such as rental agreements (for garages or buses).

evaluate the synergy measure when the combination bid constraint is binding, we set the Lagrange multipliers of the combination bid constraints to zero, which provides an upper bound on the magnitude of cost synergies. The inferred synergy measure is positive on 15 of 46 occasions only. The median cost synergy measure is negative and equals -0.24. As the measure is an upper bound on the potential cost synergy, we can conclude again that the combination of routes in a three route auction does not cost less than the sum of stand-alone costs. The synergy estimates appear qualitatively robust to changes in the assumed support of the reserve price distribution.[39]

The evidence suggests that a bidder's cost tends to increase by more than the sum of stand-alone costs. The technology appears to exhibits decreasing returns to scale. A possible explanation is that the technology of operating buses depends on the garage capacity. As the number of buses in the fleet approaches the garage capacity it becomes increasingly costly to take on additional buses or routes. Possible cost savings emerging from the sharing of spare buses and bus servicing are outweighed by cost increases due to the limited garage size.

A bound on the efficiency losses of the auction format can be inferred in the following way: First, our cost estimates give us a lower bound on the cost of the winning allocation.[40] Second, we infer costs for the remaining observed bids that did not win but that have point identified costs. We may examine whether it would have been more efficient to award one or more routes to one of these losing bids. We ignore losing bids with partially identified costs, and our comparison yields a lower bound to the efficiency loss only.[41] We illustrate our findings for the 36 auctions which consist of exactly two routes: We find that the low cost allocation differs from the observed allocation on 20 of 36 occasions with a median efficiency loss equal to 7.8% of the internal cost estimate.

We can also distinguish whether the inefficiency arises when a combination bid won or when a stand-alone bid won. We find that the main source of inefficiencies is the combination

[39] Reducing the support of the reserve prices from $[1, 1.45]$ to $[1.2, 1.45]$, has the following effects: The median cost synergy estimate equals -0.30 for two route auctions and -0.13 for three route auctions.

[40] A lower bound on the cost of the winning allocation is obtainable as the winning bid is always observed. If two stand-alone bids win, or if a combination bid wins both routes and the combination bid constraint is not binding, then the cost is point identified. If a combination bid wins, and the combination bid constraint is binding, then a lower bound to the cost of the winning allocation is identified.

[41] As described in section 3, if the upper bound of the reserve price is binding for a stand-alone bid and not for the other stand-alone bid, then the cost associated with the first stand-alone bid is not bounded from below due to the pressence of the combination bid constraint. Notice though that this concern arises for non-winning bids only, as a winning bid is below the reserve price by definition.

bid. On 13 of 20 occasions the contract was awarded to a combination bid although an award to separate bidders would have been more efficient. On the remaining 7 auctions the inefficiency arises due to asymmetries as the low cost bidder's bid exceeds the bid of another bidder although the other bidder has a higher cost. This suggests that the strategic motive for combination bidding described in section 2 may be an important element for combination bidding.

7 Conclusions

This paper has analyzed several aspects of the first price combinatorial auction: theory, identification, estimation and empirical analysis. We have analyzed bidders' incentives to use combination bids in such auctions and identified a strategic motivation for doing so that is akin to price discrimination in multi-product pricing. We have derived conditions for the model to be identified, and more importantly, proposed a method for characterizing bounds on the costs and distribution of costs when these conditions fail. We have proposed a method to estimate the distribution of bids and the probabilities of winning on the basis of data from a combinatorial auction, and have applied it to the data we collected for the auctions of London bus routes. Based on our theoretical analysis and our identification results, we have looked at empirical evidence for cost synergies in the London bus market.

All four aspects of our problem deserve further research. First, a full characterization of equilibrium in the combinatorial first price auction (including algorithms to compute the equilibrium) and further results on revenue/cost ranking of auction formats would be useful. Second, our identification results suggest that partial identification of the distribution of costs is likely to be the rule rather than the exception in multi-unit auction data sets. We need new results that allow us to answer meaningful policy questions even when the primitives are only partially identified (an example along these lines for single-unit English auctions is Haile and Tamer, 2003). Partial identification has also implications for the potential generalization of our results to affiliated settings. Existing identification results with affiliation use bidders' bids as a sufficient statistic for their private signal when identifying the probabilities of winning. Partial identification means that a bid no longer identifies a single cost. Thus generalization to affiliated settings is no longer trivial. Third, we have attributed the unexplained variations in the bid data to bidder specific private information, and assumed that there is no unobserved route and bidder heterogeneity. Our assumption is satisfied if the econometrician has access to the same information as market participants.

While this assumption is common in most of the empirical auction literature, the question whether the estimation method can be extended to accommodate unobserved heterogeneity is a challenging one.

8 Appendix A: Proofs

Lemma 1: *Consider the private value combinatorial first auction model described in Section 2, without a reserve price or with a known reserve price, and with or without combination bid constraints. Then, in any equilibrium where bidders bid on or above their costs:*

(1) Bidders are indifferent about the way in which ties that occur with positive probability are resolved,

(2) $\lim_{\mathbf{b}\to\mathbf{b}^*}\sum_{s\subseteq S}(b_s - c_s^i)G_s(\mathbf{b}) = \sum_{s\subseteq S}(b_s^* - c_s^i)G_s(\mathbf{b}^*)$ for all $\mathbf{b}^* \in \mathcal{B}_i$, such that $\mathbf{b}^*$ is bidder i's optimal bid given costs $\mathbf{c}^i$.

(3) $\eta^i(\mathcal{B}_{i3}) = 0$ *for all* i.

Proof of Lemma 1: (1) and (2): Consider bidder i. Let $z_s = \min_{t\neq s}\{\underline{b}_{S\backslash t}^{-i}+b_t, b_S, \underline{b}_S^{-i}\}-\underline{b}_{S\backslash s}^{-i}$ where $\underline{b}_s^{-i}$ were defined in (1). Bidder i's probability of winning exactly route (or route combination) $s \neq S$, G_s, given his submitted bid $\mathbf{b}$ is

$$\Pr(b_s < z_s) + \alpha\ \Pr(b_s = z_s)$$

where $\alpha \in [0,1]$ is induced by the tie-breaking rule. This probability is discontinuous with respect to b_s at $\mathbf{b}$ if z_s has a mass point, say at $\widehat{z}_s$. We need to consider three cases depending on what causes the mass point at $\widehat{z}_s$:

1. $b_s + \underline{b}_{S\backslash s}^{-i}$ ties with positive probability with $\underline{b}_{S\backslash t}^{-i} + b_t$ or b_S (but not with $\underline{b}_S^{-i}$) :

   Let $G_o(\mathbf{b})$ be the probability that bidder i does not win anything with bid $\mathbf{b}$ or, in other words, that the winning bid is $\underline{b}_S^{-i}$. By construction, $G_o(\mathbf{b}) = 1 - \sum_s G_s(\mathbf{b})$. It is continuous in b_s at $\mathbf{b}$ since, by assumption, $b_s + \underline{b}_{S\backslash s}^{-i}$ does not tie with positive probability with $\underline{b}_S^{-i}$. Thus, so is $\sum_s G_s(\mathbf{b})$. This implies that the only way in which it could be optimal to submit b_s is if $(b_s - c_s)$ is equal to the profit bidder i makes from the other allocations with which it ties. Otherwise, submitting a bid on route s slightly above or slightly below b_s is a profitable deviation. Thus, when a tie occurs, bidder i is indifferent about the way in which the tie is resolved. Moreover, his expected payoff is continuous.

2. $b_s + \underline{b}_{S\backslash s}^{-i}$ ties with positive probability with $\underline{b}_S^{-i}$ (but no other bids).

   (a) Consider first the case where the tie-breaking rule is such that $\lim_{\varepsilon\downarrow 0} G_s(b_s - \varepsilon, \mathbf{b}_{-s}) - G_s(\mathbf{b}) > 0$ (for example, if the allocation is random in case of ties). If $b_s - c_s > 0$, submitting $b_s - \varepsilon$ strictly increases bidder i's expected profit for ε small enough. If $b_s - c_s = 0$ but $G_s(b_s + \varepsilon, \mathbf{b}_{-s}) > 0$ for some $\varepsilon > 0$, submitting $b_s + \varepsilon$ is a profitable deviation. Thus the only way in which b_s could be part of an equilibrium is if $b_s - c_s = 0$ and $G_s(b_s + \varepsilon, \mathbf{b}_{-s}) = 0$ for all $\varepsilon > 0$.[42] But then expected payoff is again continuous. Moreover, bidder i is again indifferent about the way in which the tie is resolved.

(b) To complete the argument we argue that $\lim_{\varepsilon \downarrow 0} G_s(b_s - \varepsilon, \mathbf{b}_{-s}) - G_s(\mathbf{b}) = 0$ cannot happen in positive probability ties.[43] To see this, suppose first that bidder j is responsible for the mass point in $\underline{b}_S^{-i} - \underline{b}_{S\setminus s}^{-i}$. His bid on S ties with positive probability with his bid on $S\setminus s$. Thus, applying the argument in point 1 above means that his payoff from winning either allocation is the same. Given assumption 1, the set of costs satisfying this condition has zero measure, a contradiction with the fact that bidder j causes a mass in the distribution of $\underline{b}_S^{-i} - \underline{b}_{S\setminus s}^{-i}$. The same reasoning holds if bidder j is responsible for the mass in the distribution of $\underline{b}_S^{-i}$ and bidder k is responsible for the mass in the distribution of $\underline{b}_{S\setminus s}^{-i}$. Likewise if instead of a single bidder being responsible for the mass point, several bidders having a mass on subsets of S (resp. $S\setminus s$) are responsible for the mass points.

3. $b_s + \underline{b}_{S\setminus s}^{-i}$ ties with positive probability with $\underline{b}_S^{-i}$ and some other bids.

   We can combine the arguments from the two previous scenarios to conclude the bidder is indifferent about the ways ties are resolved and that expected payoff is again continuous.

Note that none of the arguments above relied on a particular structure on the bids. Thus they continue to hold when a combination bid constraint is imposed. The proof also extends to reserve prices. Reserve prices can be considered as bids submitted by the buyer, with the exception that no best response requirements is placed on them. The only place where the argument above relies on the behavior of bidder i's opponents is in part (b) of point 2, which is never relevant for reserve prices given that $\lim_{\varepsilon \downarrow 0} G_s(b_s - \varepsilon, \mathbf{b}_{-s}) - G_s(\mathbf{b}) > 0$ always for reserve prices. This establishes statements (1) and (2) of Lemma 1.

We now proceed to proving statement (3) of Lemma 1. Note first that, by Lemma 1(2), the only non differentiability points of bidders' expected payoff correspond to kinks. The rest of the proof proceeds in three steps:

[42] $b_s - c_s < 0$ is ruled out by assumption 2.

[43] This argument is only needed for proving part (1) of the Lemma. It is not needed to prove continuity of the equilibrium payoffs under the London bus routes tie-breaking rule.

Step 1: Kinks that can induce mass points in the distribution of bids.

Not all kinks are candidates to induce a mass point in the distribution of best response bids. Consider bidder i's optimization problem, $\max_{\mathbf{b}\in\mathbb{R}^{2^m-1}} \sum_{s\subseteq S}(b_s - c_s)G_s(\mathbf{b})$ for some fixed cost $\mathbf{c}$. Denote by $\frac{\partial}{\partial b_s}G_t(\mathbf{b})|_{\text{left}}$ and $\frac{\partial}{\partial b_s}G_t(\mathbf{b})|_{\text{right}}$, the left and right derivatives of G_t with respect to b_s at $\mathbf{b}$. A necessary condition for kinks to generate a mass in bidder i's distribution of best response bids is that, for some $t \subset S$ and some $\mathbf{b}$,

$$G_t(\mathbf{b}) + \sum_{s\subseteq S}(b_s - c_s)\frac{\partial}{\partial b_t}G_s(\mathbf{b})|_{\text{left}} \geq 0 \geq G_t(\mathbf{b}) + \sum_{s\subseteq S}(b_s - c_s)\frac{\partial}{\partial b_t}G_s(\mathbf{b})|_{\text{right}} \tag{15}$$

with at least one inequality strict. Intuitively, suppose (15) is satified for $\mathbf{b}$ and $\mathbf{c}$. Then it is also satisfied for $\mathbf{b}$ and $\mathbf{c}'$ in the neighborhood of $\mathbf{c}$. Bids $\mathbf{b}$ such that

$$G_t(\mathbf{b}) + \sum_{s\subseteq S}(b_s - c_s)\frac{\partial}{\partial b_t}G_s(\mathbf{b})|_{\text{left}} \leq 0 \leq G_t(\mathbf{b}) + \sum_{s\subseteq S}(b_s - c_s)\frac{\partial}{\partial b_t}G_s(\mathbf{b})|_{\text{right}}$$

with at least one inequality strict are never a best response for bidder i with cost $\mathbf{c}$.
Practically, consider $z_s = \min_{t\neq s}\{\underline{b}_{S\setminus t}^{-i} + b_t, b_S, \underline{b}_S^{-i}\} - \underline{b}_{S\setminus s}^{-i}$, the bid on s that bidder i must beat in order to win s. Kinks that are candidates to induce mass point bidding correspond to upward jumps in the density of z_s because they imply $\frac{\partial}{\partial b_s}G_s(\mathbf{b})|_{\text{left}} > \frac{\partial}{\partial b_s}G_s(\mathbf{b})|_{\text{right}}$ and $\frac{\partial}{\partial b_t}G_s(\mathbf{b})|_{\text{left}} \geq \frac{\partial}{\partial b_t}G_s(\mathbf{b})|_{\text{right}}$. Bids at downward jumps cause $\frac{\partial}{\partial b_s}G_s(\mathbf{b})|_{\text{left}} < \frac{\partial}{\partial b_s}G_s(\mathbf{b})|_{\text{right}}$ and $\frac{\partial}{\partial b_t}G_s(\mathbf{b})|_{\text{left}} \leq \frac{\partial}{\partial b_t}G_s(\mathbf{b})|_{\text{right}}$ and can thus never be best responses.

Step 2: $z_S = \min_{s\neq S}\{b_s^i + \underline{b}_{S\setminus s}^{-i}, \underline{b}_S^{-i}\}$ has no kink in its distribution that generate mass points in the distribution of best response bids by bidder i.

Fix $\mathbf{b}_{-S}^i$ and define $z_S = \min_{s\neq S}\{b_s^i + \underline{b}_{S\setminus s}^{-i}, \underline{b}_S^{-i}\}$ with support in $[\underline{z}_S, \overline{z}_S]$. Note that any bidder-bid combination with realization in $[\underline{z}_S, \overline{z}_S)$ has a strictly positive probability of winning. Thus the bids that make up for those allocations must be strictly above costs.
Suppose the distribution of z_S has a kink at $\widehat{z}_S$ that induces a mass point in the distribution of best response bids by bidder i. From step 1, this kink must be caused by an upward jump in its density at $\widehat{z}_S$. Upward jumps can happen for two reasons:

1. The random variables $b_s^i + \underline{b}_{S\setminus s}^{-i}$ and $\underline{b}_S^{-i}$ do not all have the same support and $\widehat{z}_S$ is at the lower bound of the support of one of these variables.

2. The random variables $b_s^i + \underline{b}_{S\setminus s}^{-i}$ or $\underline{b}_S^{-i}$ have a discontinuous density for values for which they have a strictly positive probability of winning.

We first consider the case where the kink happens at the boundary of the support of $\underline{b}_S^{-i}$ or because of a discontinuity of its density. Towards a contradiction suppose that, as a result, $\eta^i(\{\widetilde{\mathbf{b}}^i \in \mathcal{B}_i : \widetilde{b}_S^i = \widehat{z}_S\}) > 0$. We consider two subcases:

1. <u>There exists $j \neq i$ such that the density of b_S^j has an upward jump at $\widehat{z}_S$</u> . Bidder j's bid ties with strictly positive probability with b_S^i. From lemma 1(1), it must be that $b_S^j = c_S^j$. This contradicts optimality because this bid wins with strictly positive probability. We conclude that the kink cannot be caused by a discontinuity in the density of b_S^j after all.

2. <u>There exist $j, k \neq i$ such that the density of $b_{S\backslash t}^j + b_t^k$ is discontinuous at $\widehat{z}_S$ when $b_{S\backslash t}^j + b_t^k = \underline{b}_S^{-i} = z_S$</u>: (note: if the kink is made of more than two bids, the same reasoning applies). Because $b_{S\backslash t}^j + b_t^k$ is the sum of two independently distributed random variables, its density is continuous on its support and vanishes to zero at the boundary of its support, ruling out a kink unless one of $b_{S\backslash t}^j$ or b_t^k has a mass point. Suppose, without loss of generality, that b_t^k has a mass point at x. Let $b_{S\backslash t,L}^j$ denote the lower bound of the support of $b_{S\backslash t}^j$. We have $\widehat{z}_S = x + b_{S\backslash t,L}^j$. Consider bidder j. His bid $b_{S\backslash t,L}^j$ ties with positive probability with b_S^i. By lemma 1(1), $b_{S\backslash t,L}^j = c_{S\backslash t}^j$. This contradicts optimality because this bid wins with strictly positive probability. We conclude that the kink cannot be caused by a kink in the distribution of $b_{S\backslash t}^j + b_t^k$.

We next consider the case where the kink happens at the lower bound of the support of $b_s^i + \underline{b}_{S\backslash s}^{-i}$ for some s, or at an upward jump in the density of $b_s^i + \underline{b}_{S\backslash s}^{-i}$. Let $x = \widehat{z}_S - b_s^i$. Suppose that the resulting kink implies that, at equilibrium, $\eta^i(\{\widetilde{\mathbf{b}}^i \in \mathcal{B}_i : \widetilde{b}_S^i = \widetilde{b}_s^i + x$, $G_S^i(\widetilde{\mathbf{b}}^i) > 0$ and $G_s^i(\widetilde{\mathbf{b}}^i) > 0\}) > 0$. We consider two subcases.

1. <u>There exists $j \neq i$ such that the density of $b_{S\backslash s}^j$ has an upward jump at x</u>. Bidder j's bid ties with strictly positive probability with bidder i's bid on S. Thus from Lemma 1(1), it must be that $b_{S\backslash s}^j = c_{S\backslash s}^j$. This contradicts optimality because this bid wins with strictly positive probability. We conclude that the kink at $\widehat{z}_S$ cannot have been caused by a discontinuity in the density of $b_{S\backslash s}^j$.

2. <u>There exists $k, j \neq i$, such that the density of $b_w^k + b_{w'}^j$ ($w \cap w' = \varnothing$ and $w \cup w' = S\backslash s$), is discontinuous at x</u>: Because is the sum of two independently distributed random variables, it must be that one of them has a mass point. Suppose b_w^k has a mass point at y. Thus $x = y + b_{w',L}^j$ ($b_{w',L}^j$ is the lower bound of the support of $b_{w'}^j$). Consider

bidder j : his bid on w' ties with positive probability with b^i_S. Thus $b^j_{w',L} = c^j_{w'}$. This contradicts optimality because this bid has a strictly positive probability of winning.

We conclude that z_S has no kink that generates mass points in the distribution of bids by bidder i.

Step 3: $z_t = \min_{s \neq t}\{b^i_s + \underline{b}^{-i}_{S\backslash s}, \underline{b}^{-i}_S, b^i_S\} - \underline{b}^{-i}_{S\backslash t}$, $t \neq S$, has no kink that generates a mass point in the distribution of best response bids by bidder i.

Random variable z_t differs from z_S in two respects. First, the presence of b^i_S in the definition of z_t means that $\min_{s \neq t}\{b^i_s + \underline{b}^{-i}_{S\backslash s}, \underline{b}^{-i}_S, b^i_S\}$ has necessarily a mass point at its upper bound, b^i_S. Second, the fact that z_t is the difference between two random variables means that a necessary condition for z_t to have a kink is that one of $\min_{s \neq t}\{b^i_s + \underline{b}^{-i}_{S\backslash s}, \underline{b}^{-i}_S, b^i_S\}$ or $\underline{b}^{-i}_{S\backslash t}$ has a mass point and that the other does not have a vanishing density at its boundaries (unless $\min_{s \neq t}\{b^i_s + \underline{b}^{-i}_{S\backslash s}, \underline{b}^{-i}_S, b^i_S\}$ and $\underline{b}^{-i}_{S\backslash t}$ are perfectly correlated which could only be the case if $N = 2$ but, in this case, z_S has also a kink and thus this case is covered by the previous one). We distinguish four cases:

1. The kink in z_t is due to the mass point in the distribution of $\min_{s \neq t}\{b^i_s + \underline{b}^{-i}_{S\backslash s}, \underline{b}^{-i}_S, b^i_S\}$ at b^i_S :

   This creates an upward jump in the density of z_t at $b^i_S - \underline{b}^{-i}_{S\backslash t,H}$, where $\underline{b}^{-i}_{S\backslash t,H}$ is the lower bound to $\underline{b}^{-i}_{S\backslash t}$. Suppose that as a result $\eta^i(\{\widetilde{\mathbf{b}}^i \in \mathcal{B}_i : \widetilde{b}^i_t + \underline{b}^{-i}_{S\backslash t,H} = \widetilde{b}^i_S$, $G^i_t(\widetilde{\mathbf{b}}^i) > 0\}) > 0$. We consider two subcases:

   (a) There exists $j \neq i$ such that $b^j_{S\backslash t} = \underline{b}^{-i}_{S\backslash t,H}$. Bidder j's bid ties with positive probability with $\underline{b}^{-j}_S$, implying that $b^j_{S\backslash t} = c_{S\backslash t}$ (Lemma 1(1)). But this contradicts optimality given that this bid has a strictly positive probability of winning.

   (b) There exists $j, k \neq i$ such that $b^j_{w,H} + b^k_{w',H} = \underline{b}^{-i}_{S\backslash t,H}$ ($w \cap w' = \varnothing$ and $w \cup w' = S\backslash t$). Because by assumption $\underline{b}^{-i}_{S\backslash t}$ has a non vanishing density at $\underline{b}^{-i}_{S\backslash t,H}$, it must be that j or k has a mass on their bids. Suppose j has a mass at $b^j_{w,H}$. Consider bidder k. His bid ties with strictly positive probability with b^i_S, implying that $b^j_w = c_w$ (Lemma 1(1)). But this contradicts optimality given that this bid has a strictly positive probability of winning.

2. The kink in z_t is due to a mass point in the distribution of $\underline{b}^{-i}_{S\backslash s}$ at x when $G_s > 0$. This creates an upward jump in the density of z_t at $b^i_s + x - \underline{b}^{-i}_{S\backslash t,H}$ ($\underline{b}^{-i}_{S\backslash t,H}$ is the upper bound to $\underline{b}^{-i}_{S\backslash t}$). Suppose that, as a result, $\eta^i(\{\widetilde{\mathbf{b}}^i : \widetilde{b}^i_t - \widetilde{b}^i_s = x - \underline{b}^{-i}_{S\backslash t,H}$, $G_s(\widetilde{\mathbf{b}}^i)$, $G_t(\widetilde{\mathbf{b}}^i) > 0\}) > 0$ holds. We consider two subcases.

(a) There exists $j \neq i$ such that $b^j_{S\backslash t} = \underline{b}^{-i}_{S\backslash t,H}$. If j is not responsible for the mass at x, his bid ties with positive probability with $\underline{b}^{-j}_S$, implying that $b^j_{S\backslash t} = c_{S\backslash t}$ (Lemma 1(1)). This contradicts optimality given that this bid has a strictly positive probability of winning. If, instead j is responsible for the mass at x, then his bids on $S\backslash s$ and $S\backslash t$ tie with positive probability. By Lemma 1(1), he must be indifferent between the profits he makes from both bids. A contradiction with the mass that he places at x.

(b) There exists $j, k \neq i$ such that $b^j_{w,H} + b^k_{w',H} = \underline{b}^{-i}_{S\backslash t,H}$ ($w \cap w' = \varnothing$ and $w \cup w' = S\backslash t$). Because by assumption $\underline{b}^{-i}_{S\backslash t}$ has a non vanishing density at $\underline{b}^{-i}_{S\backslash t,H}$, it must be that j or k has a mass on their bids. Suppose j has a mass at $b^j_{w,H}$. Consider bidder k. Suppose bidder k is not responsible for the mass in the distribution of $\underline{b}^{-i}_{S\backslash s}$. Then, his bid ties with positive probability with $\underline{b}^{-k}_S$ implying $c^k_{w'} = b^k_{w',H}$.A contradiction with the fact that this bid wins with strictly positive probability. If instead bidder k is responsible for the mass in $\underline{b}^{-i}_{S\backslash s}$ Then his bids on $S\backslash s$ and w' tie with positive probability. By lemma 1(1), he must be indifferent between the profits he makes from both bids. A contradiction with the mass that he places a mass at x.

3. The kink in z_t is due to a mass point in the distribution of $\underline{b}^{-i}_S$ at x. This creates an upward jump in the density of z_t at $x - \underline{b}^{-i}_{S\backslash t,H}$. Suppose that as result, $\eta(\widetilde{b}^i : \widetilde{b}^i_t = x - \underline{b}^{-i}_{S\backslash t,H}, G_t > 0) > 0$. Same arguments as before.

4. The kink in z_t is due to a mass point in the distribution of $\underline{b}^{-i}_{S\backslash t}$: this case is identical to case 2 above, once we consider z_s instead of z_t.

Steps 2 and 3 imply that there is no kink in bidder i's objective function that induces a mass point in the distribution of his best response bids. Thus $\mu(\mathcal{B}_{i3}) = 0$ implies that $\eta^i(\mathcal{B}_{i3})$. This concludes the proof for statement (3). Q.E.D.

Lemma 3: *Consider matrix* $\mathbf{\nabla G}(\mathbf{b}^i)$ *(with elements defined by* $\nabla G_{t,s}(\mathbf{b}^i) = \frac{\partial}{\partial b_t} G_s(\mathbf{b}^i)$*) evaluated at any optimal bid vector* $\mathbf{b}^i$ *by bidder* i*. Then:*

(1) $\frac{\partial}{\partial b_t} G_t(\mathbf{b}^i) \leq 0$ *for all* t*, and strictly so if* $G_t(\mathbf{b}^i) > 0$*.*

(2) $\frac{\partial}{\partial b_t} G_s(\mathbf{b}^i) \geq 0$ *for all* $t \neq s$*.*

(3) $\sum_s \frac{\partial}{\partial b_t} G_s(\mathbf{b}^i) \leq 0$ *for all* t*, and strictly so for some* t *if there exists* $s \subseteq S$ *such that* $G_s(\mathbf{b}^i) > 0$*.*

(4) $\mathbf{\nabla G}(\mathbf{b}^i)$ *is invertible if* $G_s(\mathbf{b}^i) > 0$ *for all* s*.*

(5) The determinant of any submatrix made from removing some rows and the corresponding columns of ∇G has sign (-1)^r where r is the number of remaining rows/columns if all bid components in $\mathbf{b}^i$ except those removed have a strictly positive probability of winning.

Proof of Lemma 3: (1), (2) and (3): Given m routes, there are 2^m possible allocations between bidder i and his opponents: Either $b_s^i + \underline{b}_{S\backslash s}^{-i}$ generates the cheapest allocation, in which case bidder i wins exactly contract s, or $\underline{b}_S^{-i}$ yields the lowest cost, in which case bidder i does not win anything.

Any increase in b_t^i makes the allocation that corresponds to $b_t^i + \underline{b}_{S\backslash t}^{-i}$ more expensive relative to the other ones, but it otherwise does not affect the relative ranking of $b_s^i + \underline{b}_{S\backslash s}^{-i}$, $s \neq t$, and $\underline{b}_S^{-i}$. Hence the probability that any of these competing allocations wins cannot decrease: $\frac{\partial}{\partial b_t} G_s \geq 0$ for $s \neq t$. Likewise, the probability that allocation $b_t^i + \underline{b}_{S\backslash t}^{-i}$ wins cannot increase: $\frac{\partial}{\partial b_t} G_t \leq 0$. Finally, $\sum_s \frac{\partial}{\partial b_t} G_s \leq 0$ for all t (since $\sum_s G_s$ corresponds to one minus the probability that $\underline{b}_S^{-i}$ wins).

If b_t^i has a strictly positive probability of winning, i.e. $G_t > 0$, $\frac{\partial}{\partial b_t} G_t$ must be strictly negative, for otherwise increasing b_t^i by epsilon would make bidder i strictly better off (given the previous argument, raising b_t^i does not hurt the expected profit bidder i makes from his other bids), a contradiction with the fact that $\mathbf{b}^i$ is optimal.

We now show that $\sum_s \frac{\partial}{\partial b_t} G_s < 0$ for some t when $\mathbf{b}^i$ contains at least one bid that has a positive probability of winning. Towards a contradiction, suppose that $\sum_s \frac{\partial}{\partial b_t} G_s = 0$ for all t. This means that the support of $\min_{s \subseteq S}\{b_s^i + \underline{b}_{S\backslash s}^{-i}\}$ is distinct from the support of $\underline{b}_S^{-i}$ (none of bidder i's bids compete with $\underline{b}_S^{-i}$).[44] Since one of bidder i's bids wins sometimes:[45]

$$\min_{s \subseteq S}\{b_s^i + \underline{b}_{S\backslash s}^{-i}\} \leq \underline{b}_S^{-i}, \text{ for all realizations of } \underline{\mathbf{b}}^{-i} \tag{16}$$

Since $\mathbf{b}^i$ is optimal

$$\max_{\underline{\mathbf{b}}^{-i}}\{\min_{s \subseteq S}\{b_s^i + \underline{\mathbf{b}}_{S\backslash s}^{-i} - \underline{\mathbf{b}}_S^{-i}\}\} = 0$$

(for otherwise increasing all bids by epsilon would be a profitable deviation for bidder i). But this means that $\underline{\mathbf{b}}_S^{-i}$ competes at least with one of bidder i's bids. Hence $\sum_s \frac{\partial}{\partial b_t} G_s < 0$ for at least one t.

(4) We introduce the notation "contract 0" with the convention that $b_0^i = 0$, $S\backslash\{0\} = S$, and say that bidder i wins contract 0 when $\underline{\mathbf{b}}_S^{-i}$ corresponds to the winning allocation. With this definition, $G_0 = 1 - \sum_s G_s$.

[44] The random variables here are $\underline{\mathbf{b}}^{-i}$; $\mathbf{b}^i$ is fixed.

[45] The expression in (16) implicitly assumes that the supports of $\min_{s \subseteq S}\{b_s^i + \underline{b}_{S\backslash s}^{-i}\}$ and $\underline{b}_S^{-i}$ are convex. It is straightforward to adapt the argument to non convex supports.

The following definition will be useful:

Definition: $\Omega \subseteq S \cup \{0\}$ forms a connected chain of substitutes if, for all s and $s' \subseteq \Omega$ ($s \neq s'$ and $s' \neq 0$), either $\frac{\partial}{\partial b_{s'}} G_s > 0$ or there exist $w_1, ..., w_n \subseteq \Omega$ such that $\frac{\partial}{\partial b_{w_1}} G_s^{w_1} > 0, \frac{\partial}{\partial b_{w_2}} G_{w_1} > 0, ..., \frac{\partial}{\partial b_{s'}} G_{w_n} > 0$.

Claim 1: If $G_s(\mathbf{b}^i) > 0$ for all $s \subseteq S$, then $\Omega = S \cup \{0\}$ forms a connected chain of substitutes.

Proof: By property (1), $\frac{\partial}{\partial b_s} G_s < 0$, so any contract $s \subseteq S$ must be connected with at least one other contract. By property (3), contract 0 is connected to at least one other contract. Now, if two contracts in Ω are not connected, they must exist at least two disjoint sets of contracts in Ω, with no contract in the first set connected with a contract in the other set. We prove that if $G_s(\mathbf{b}^i) > 0$ for all $s \subseteq S$, then Ω forms a connected chain of substitutes. Towards a contradiction, suppose that set $\{s, t\}$ and the rest form two disjoint sets of contracts (the focus on a set of two contracts is without loss of generality). Consider the following random variables, $\min\{b_t^i + \underline{b}_{S\backslash t}^{-i}, b_s^i + \underline{b}_{S\backslash s}^{-i}\}$ and $\min_{w \neq t,s}\{b_w^i + \underline{b}_{S\backslash w}^{-i}\}$. Since all bids have a positive probability of winning, sometimes $\min\{b_t^i + \underline{b}_{S\backslash t}^{-i}, b_s^i + \underline{b}_{S\backslash s}^{-i}\} \leq \min_{w \subseteq \Omega, w \neq t,s}\{b_w^i + \underline{b}_{S\backslash w}^{-i}\}$ (bidder i wins contract s or t) and sometimes $\min\{b_t^i + \underline{b}_{S\backslash t}^{-i}, b_s^i + \underline{b}_{S\backslash s}^{-i}\} \geq \min_{w \subseteq \Omega, w \neq t,s}\{b_w^i + \underline{b}_{S\backslash w}^{-i}\}$. Hence $\min\{b_t^i + \underline{b}_{S\backslash t}^{-i}, b_s^i + \underline{b}_{S\backslash s}^{-i}\} = \min_{w \subseteq \Omega, w \neq t,s}\{b_w^i + \underline{b}_{S\backslash w}^{-i}\}$ must happen for some realization of $\underline{\mathbf{b}}^{-i}$ given that $\mathbf{b}^i$ is optimal (if those supports were disjoint there would be a scope for a profitable deviation). Therefore, s or t must compete directly with some w in the other set, i.e. $\frac{\partial}{\partial b_t} G_w$ or $\frac{\partial}{\partial b_s} G_w > 0$. A contradiction.

We can now prove that $\det \nabla \mathbf{G} < 0$ (so that $\nabla \mathbf{G}$ is invertible). The proof is by induction. Property (3) holds strictly for at least one contract. We relabel the rows and columns of matrix $\nabla \mathbf{G}$ such that the sum of the elements in the first row is strictly negative (this does not change the value of the determinant):

$$\sum_s \frac{\partial}{\partial b_1} G_s < 0 \tag{17}$$

Consider the linear transformation L_1 on the columns of $\nabla \mathbf{G}$ that adds to column $s \neq 1$, α_{1s} times column 1 such that $\frac{\partial}{\partial b_1} G_s + \alpha_{1s} \frac{\partial}{\partial b_1} G_1 = 0$ for $s \neq 1$ (notice, $\alpha_{1s} \geq 0$ and $\sum \alpha_{1s} < 1$ given (17)). This leaves the first row of matrix $\nabla \mathbf{G}$ with all zeros except in the first position. Denote the resulting matrix by $L_1 \nabla \mathbf{G}$ and let $[L_1 \nabla \mathbf{G}]$ be matrix $L_1 \nabla \mathbf{G}$ from which the first row and the first column have been removed. Since determinants are invariant to linear transformations, $\det \nabla \mathbf{G} = \det L_1 \nabla \mathbf{G} = \frac{\partial}{\partial b_1} G_1 \det[L_1 \nabla \mathbf{G}]$.

We claim that the resulting 2^m x 2^m matrix $[L_1 \nabla \mathbf{G}]$ satisfies properties (1) to (3) of the original matrix, including the strict inequalities. Property (1): The diagonal elements of matrix $[L_1 \nabla \mathbf{G}]$ are equal to $\frac{\partial}{\partial b_s} G_s + \alpha_{1s} \frac{\partial}{\partial b_s} G_1$. Since the $\frac{\partial}{\partial b_s} G_t$ elements satisfy properties (1) to (3) and $\alpha_{1i} < 1$, we have $\frac{\partial}{\partial b_s} G_s + \alpha_{1s} \frac{\partial}{\partial b_s} G_1 < 0$. Property (2): The off-diagonal of the new matrix are equal to $\frac{\partial}{\partial b_s} G_t + \alpha_{1t} \frac{\partial}{\partial b_s} G_1 \geq 0$ since it is a sum of positive elements. Property (3): The sum of the row elements of the $[L_1 \nabla \mathbf{G}]$ matrix is equal to $\sum_{s \neq 1} \frac{\partial}{\partial b_t} G_s + \frac{\partial}{\partial b_t} G_1 \sum_{s \neq 1} \alpha_{1s} \leq 0$ since $\sum \frac{\partial}{\partial b_t} G_s \leq 0$ and $\sum_{s \neq 1} \alpha_{1s} < 1$. To show that this inequality holds strictly for at least one row of the new matrix $[L_1 \nabla \mathbf{G}]$, we need to consider two cases. First, if any of the elements $\frac{\partial}{\partial b_s} G_1$ of the first column of the original matrix was strictly positive, then since $\sum_{s \neq 1} \alpha_{1s} < 1$, there exists a row in the new matrix such that condition (3) holds strictly. If $\frac{\partial}{\partial b_s} G_1 = 0$ for all $s \neq 1$, contract 1 is directly connected only to contract 0. But then by claim 1, it must be that one of the remaining contracts, say t, is also connected to contract 0. This means that $\sum_s \frac{\partial}{\partial b_t} G_s < 0$ for that contract t in the original matrix, and in the new matrix.

Repeating the argument leads to $sign(\det \nabla \mathbf{G}) = sign(-1)^{2^m - 1} < 0$.

To prove the last part of the claim we show that any submatrix made from $\nabla \mathbf{G}$ by removing some rows and the corresponding columns has the same properties (1) to (3), including the strict inequalities. The proof then proceeds as before. Q.E.D.

Theorem 2: *Consider the combinatorial first price auction model with reserve prices and, possibly, a combination bid constraint.*

(1) ***Identification of costs.*** *Consider any equilibrium bid vector* $\mathbf{b}^i \in \mathcal{B}_{i1} \cup \mathcal{B}_{i2}$.

(i) Suppose $G_s(\mathbf{b}^i) > 0$ *for some* $s \subseteq S$*, then upper and lower bounds to* c_s^i are identified following (12) and (13). *In particular, if* $G_s(\mathbf{b}^i) > 0$ *for all* $s \subseteq S$ *and* $\mathbf{b}^i$ *satisfies strictly all combination bid constraints (if any applies), then* $\mathbf{c}^i$ *is point identified.*

(ii) Suppose $G_s(\mathbf{b}^i) = 0$ *for some* $s \subseteq S$ *and the auction does not have a combination bid constraint. Then a lower bound to* c_s^i, $\underline{c}_s^i$, *is identified following (9). If* $\overline{b}_s^i(\mathbf{b}_{-s}^i) = R_s$ *then* $\underline{c}_s^i = R_s$.

(iii) Suppose $G_s(\mathbf{b}^i) = 0$ *for some* $s \subseteq S$ *and the auction has a combination bid constraint. Then a lower bound to* c_s^i, $\underline{c}_s^i$, *is identified following (9) if* $G_t(\mathbf{b}^i) > 0$ or $G_t(\overline{b}_s^i(\mathbf{b}_{-s}^i), \mathbf{b}_{-s}^i) = 0$ *for all* $t \supset s$. No bound is identified otherwise.

(2) ***(Partial) Identification of the distributions of costs.*** *For all* i, let $\phi^i(.) : \mathcal{B}_{i1} \cup \mathcal{B}_{i2} \rightrightarrows \mathbb{R}^{2^m - 1}$ *define the mapping from bidder* i's *equilibrium bids to the set of costs (obtained from the application of (1)) that are consistent with them. Any* $F_i \in \mathcal{F}$ such that $F^i(\mathbf{c}) = \eta^i(\{\mathbf{b}^i \in \mathcal{B}_{i1} \cup \mathcal{B}_{i2} : \widetilde{\phi}^i(\mathbf{b}^i) \leq \mathbf{c}\})$ for some selection $\widetilde{\phi}^i$ of ϕ^i, is consistent with the observed equilibrium.

Proof of Theorem 2:

Part (1). (i) and the first part of (ii) follow directly from the arguments in the main text. We prove that if $\overline{b}_s^i(b_{-s}^i) = R_s$ then $\underline{c}_s^i = R_s$ when there is no combination bid constraint. When the reserve price binds, $\lim_{\widehat{b}_s \uparrow R_s} G_s(\widehat{b}_s, \mathbf{b}_{-s}^i) > 0$. From Lemma 1(1), we can consider without loss of generality that $G_s(R_s, \mathbf{b}_{-s}^i) = 0$. Since bidder i did not submit a bid on contract s below the reserve price, the following must hold for all $\widehat{b}_s < R_s$:

$$\sum_{t \neq s}(b_t^i - c_t)G_t(\widehat{b}_s, \mathbf{b}_{-s}^i) + (\widehat{b}_s - c_s)G_s(\widehat{b}_s, \mathbf{b}_{-s}^i) \leq \sum_{t \neq s}(b_t^i - c_t)G_t(R_s, \mathbf{b}_{-s}^i) \tag{18}$$

Because $\lim_{\widehat{b}_s \uparrow R_s} G_t(\widehat{b}_s, \mathbf{b}_{-s}^i) = G_t(R_s, \mathbf{b}_{-s}^i)$, for all $t \neq s$, the claim follows directly from the fact that $\lim_{\widehat{b}_s \uparrow R_s} G_s(\widehat{b}_s, \mathbf{b}_{-s}^i) > 0$. This completes the proof for (ii).

We now consider the case where there is a combination bid constraint. The difference with the previous argument is that a bid on contract s can define a bid on contract $w \supset s$. Thus a bidder may not want to submit a bid on a contract s for fear that doing so would generate, via the combination bid constraint, a bid on contract w that has a positive probability of winning. Formally, the equivalent of expression (9) is now (s is a contract such that $G_s(\mathbf{b}^i) = 0$, $G_t(\mathbf{b}^i) > 0$ for $t \in K$)

$$\begin{aligned}
&\lim_{b_s \uparrow \overline{b}_s^i} G_s(b_s, \mathbf{b}_{-s}^i) + \sum_{t \in K}(b_t^i - c_t^i)\frac{\partial}{\partial b_s}G_t(\overline{b}_s^i, \mathbf{b}_{-s}^i) + (\overline{b}_s^i - c_s^i)\frac{\partial}{\partial b_s}G_s(\overline{b}_s^i, \mathbf{b}_{-s}^i) \\
&+ \sum_{w \supset s, G_w(\mathbf{b}^i)=0 \text{ but } G_w(\overline{b}_s^i, \mathbf{b}_{-s}^i)>0} (\overline{b}_s^i + b_{w \setminus s}^i - c_w^i)\frac{\partial}{\partial b_s}G_w(\overline{b}_s^i, \mathbf{b}_{-s}^i) \ \geq 0
\end{aligned}$$

Unlike before we now have several unknows in this expression, c_s^i as well as the c_w^i's if there exists $w \supset s, G_w(\mathbf{b}^i) = 0$ but $G_w(\overline{b}_s^i, \mathbf{b}_{-s}^i) > 0$. Proceeding by iteration, one could first identify cost ranges for the c_w^i's: $c_w^i \geq \underline{c}_w^i$. However, with binding combination bid constraints, the left derivative $\frac{\partial}{\partial b_s}G_w(\overline{b}_s^i, \mathbf{b}_{-s}^i)$ is negative thus the bounds on c_w^i are useless to pin down a bound on c_s^i. If there is no w such that $w \supset s, G_w(\mathbf{b}^i) = 0$ but $G_w(\overline{b}_s^i, \mathbf{b}_{-s}^i) > 0$, then the last term disappears and we are back to (9).

Part (2): The arguments in part (1) covered all the bids in $\mathcal{B}_{i1}$ and $\mathcal{B}_{i2}$ which is enough given that by Lemma 1(3), bids in $\mathcal{B}_{i3}$ are submitted by a zero measure of costs. The rest of the claim follows. Q.E.D.

Theorem 3: *(1) For any bundle s that wins with strictly positive probability, c_s^i depends positively on the value of $\lambda_{s=t \cup w}$, and negatively on $\lambda_{t=s \cup w}$ for all s, t and w.*

(2) Consider any t and w such that $t \cap w = \emptyset$. *If the combination bid constraint for these contracts is the only binding combination bid constraint at the optimum, an upper bound to*

the synergy involved between these two contracts is given by the solution $c_t + c_w - c_{t\cup w}$ *of the system in (11) when the Lagrangian multiplier* $\lambda_{\{t\cup w\}=t\cup w}$ *is set equal to zero.*

Proof of Theorem 3:

Preliminaries: The proof uses the following properties of determinants: (1) Determinants are invariant to linear transformations of rows or columns, (2) permuting rows (or columns) just changes the sign of the determinant, (3)

$$\det\begin{bmatrix} a_{11}+b_{11} & .. & a_{1N} \\ .. & & .. \\ a_{N1}+b_{N1} & .. & a_{NN}\end{bmatrix} = \det\begin{bmatrix} a_{11} & .. & a_{1N} \\ .. & & .. \\ a_{N1} & .. & a_{NN}\end{bmatrix} + \det\begin{bmatrix} b_{11} & .. & a_{1N} \\ .. & & .. \\ b_{N1} & .. & a_{NN}\end{bmatrix}, \text{ and (4)}$$

the multiplication of any row or column by a constant, multiplies the value of the determinant by that constant.

Proof of Part (1): Consider any equilibrium bid vector $\mathbf{b}^i$. Let $K \subset 2^S$ be the set of all contracts which $\mathbf{b}^i$ wins with strictly positive probability and let $\mathbf{b}^i_K$ denote the bids in $\mathbf{b}^i$ restricted to K (similarly, $\mathbf{c}^i_K$ denote the cost vector associated with K). >From the derivation in the main text,

$$(\mathbf{b}^i_K - \mathbf{c}^i_K)\nabla\widehat{\mathbf{G}}(\mathbf{b}^i_K) = \mathbf{D}(\mathbf{b}^i_K, \boldsymbol{\lambda})$$

where

$$\mathbf{D}(\mathbf{b}^i_K, \boldsymbol{\lambda}) = -\widehat{\mathbf{G}}(\mathbf{b}^i_K) + \sum_r \sum_{\substack{t,w\subset r,\ t\cap w=\emptyset, \\ |t|\leq|w|}} \lambda_{r=t\cup w}\mathbf{I}_{r=t\cup w}$$

with $\mathbf{I}_{r=t\cup w}$, the $|K|$ x 1 vector with entry 1 in the row corresponding to contract r and entry -1 in the rows corresponding to contracts t and w.

Let $\mathbf{A}_s\mathbf{B}$ denote matrix $\mathbf{A}$ whose column corresponding to contract s has been replaced by vector $\mathbf{B}$. Cramer's rule together with properties (3) and (4) of determinants imply that

$$\begin{aligned} b^i_s - c^i_s &= \frac{1}{\det \boldsymbol{\nabla}\widehat{\mathbf{G}}} \det \boldsymbol{\nabla}\widehat{\mathbf{G}}_s\mathbf{D}(\mathbf{b}^i_K; \boldsymbol{\lambda}) \\ &= \frac{1}{\det \boldsymbol{\nabla}\widehat{\mathbf{G}}}[-\det \boldsymbol{\nabla}\widehat{\mathbf{G}}_s\widehat{\mathbf{G}} + \sum_{r\in K} \sum_{\substack{t,w\subset r,\ t\cap w=\emptyset, \\ |t|\leq|w|}} \lambda_{r=t\cup w} \det \boldsymbol{\nabla}\widehat{G}_s\mathbf{I}_{r=t\cup w}] \end{aligned}$$

Step 1: $\det \boldsymbol{\nabla}\widehat{\mathbf{G}}_r\mathbf{I}_{r=t\cup w} = 0$ or, if not, $\text{sign}\big(\det \boldsymbol{\nabla}\widehat{\mathbf{G}}_r\mathbf{I}_{r=t\cup w}\big) = \text{sign}(\text{-}1^{|K|-1})$.

Proof: $\boldsymbol{\nabla}\widehat{\mathbf{G}}_r\mathbf{I}_{r=t\cup w}$ is the $|K|$ by $|K|$ matrix $\boldsymbol{\nabla}\widehat{\mathbf{G}}$ whose column r is replaced by a column with +1 at row r, and -1 at rows t and w. Define L_r as the operator that adds the values associated with row r to rows t and w so that $L_r\boldsymbol{\nabla}\mathbf{G}_r\mathbf{I}_{r=t\cup w}$ becomes a matrix with a zero column at position r except for the "1" entry at row r. Define $\mathbf{M}$ as the $|K|-1$ by $|K|-1$

matrix made of matrix $L_r \boldsymbol{\nabla} \widehat{G}_r \mathbf{I}_{r=t\cup w}$ from which the row and the column corresponding to contract r have been removed. By construction and property (1) of determinants, $\det \mathbf{M} = \det L_r \boldsymbol{\nabla} \widehat{\mathbf{G}}_r \mathbf{I}_{r=t\cup w} = \det \boldsymbol{\nabla} \widehat{\mathbf{G}}_r \mathbf{I}_{_{r=t\cup w}}$.

Matrix $\mathbf{M}$ has the following properties: (a) All diagonal elements are negative. This follows from Lemma 3(1) for the rows associated with contracts $\neq t$ and w. The diagonal elements for rows t and w are $\frac{\partial}{\partial b_r} G_t + \frac{\partial}{\partial b_t} G_t$ and $\frac{\partial}{\partial b_r} G_w + \frac{\partial}{\partial b_w} G_w$ respectively, which are also negative by Lemma 3(3). (b) All off-diagonal elements are positive. Again this follows directly from Lemma 3(2) in the case of rows $\neq t, w$. For rows t and w, the diagonal elements are $\frac{\partial}{\partial b_r} G_u + \frac{\partial}{\partial b_t} G_u$, $u \neq r, t$, and $\frac{\partial}{\partial b_r} G_u + \frac{\partial}{\partial b_w} G_u$, $u \neq r, w$, which are positive by Lemma 3(2). (c) The sum of the column elements are negative for every column. Consider for example column u. The sum of its elements are $\sum_{x\neq r,t,w} \frac{\partial}{\partial b_x} G_u + (\frac{\partial}{\partial b_r} G_u + \frac{\partial}{\partial b_t} G_u) + (\frac{\partial}{\partial b_r} G_u + \frac{\partial}{\partial b_w} G_u)$, in other words, the change in the probability of winning u when all bids are marginally increased (with the constraint that $b_r = b_t + b_w$). Clearly, this cannot increase the probability that the bidder wins contract u.

We can now apply a similar induction technique as used in Lemma 3 to prove that, if $\det \mathbf{M}$ is not zero, $\text{sign}(\det \mathbf{M}) = \text{sign}(-1^{|K|-1})$. The two differences relative to the proof in Lemma 3 are that we work on the columns instead of the rows and that all inequalities are weak.

Define operator P^1 which operates on the rows of matrix $\mathbf{M}$ and adds to row $u \neq 1$, α_{u1} such that $M_{u1} + \alpha_{u1} M_{11} = 0$ ($\alpha_{u1} \geq 0$ and $\sum_u \alpha_{u1} \leq 1$).[46] It is straightforward to check that the resulting matrix, $P^1\mathbf{M}$, satisfies properties (a) through (c). Therefore we can reiterate this process until the end or until the first column of the resulting matrix is all zeroes. This proves the claim.

Step 2: $\det \boldsymbol{\nabla} \widehat{\mathbf{G}}_t \mathbf{I}_{_{r=t\cup w}} = 0$ or, if not, $\text{sign}\big(\det \boldsymbol{\nabla} \widehat{\mathbf{G}}_t \mathbf{I}_{_{r=t\cup w}}\big) = -\ \text{sign}(\text{-}1^{|K|-1})$.

Proof: Matrix $\boldsymbol{\nabla} \widehat{\mathbf{G}}_t \mathbf{I}_{_{r=t\cup w}}$ corresponds to matrix $\boldsymbol{\nabla} \widehat{\mathbf{G}}$, whose column t has been replaced by a column with $+1$ at position r and -1 at positions t and w. In particular, note that position (t,t) is equal to -1. Define the operator L_t that adds row r to row w and adds row t to row r so that only a -1 at position t remains in column t. Define $\mathbf{M}$ as the $|K| - 1$ by $|K| - 1$ matrix made of matrix $L_t \boldsymbol{\nabla} \widehat{G}_t \mathbf{I}_{r=t\cup w}$ from which the row and the column corresponding to contract t have been removed. By construction and property (1) of determinants, $\det \mathbf{M} = -\det L_t \boldsymbol{\nabla} \widehat{\mathbf{G}}_t \mathbf{I}_{r=t\cup w} = -\det \boldsymbol{\nabla} \widehat{\mathbf{G}}_t \mathbf{I}_{_{r=t\cup w}}$.

Matrix $\mathbf{M}$ has the following properties: (a) All diagonal elements are negative. This follows from Lemma 3(1) for the rows associated with contracts $\neq r$ and w. The diagonal elements

[46] If $M_{11} = 0$, the first column contains only zero entries given properties (b) and (c) so we are done.

for rows r and w are $\frac{\partial}{\partial b_r}G_r + \frac{\partial}{\partial b_t}G_r$ and $\frac{\partial}{\partial b_r}G_w + \frac{\partial}{\partial b_w}G_w$ respectively, which are also negative by Lemma 3(3).

(b) All off-diagonal elements are positive, except for the entry in row w and column r. Again this follows directly from Lemma 3(2) in the case of rows $\neq r, w$. The off-diagonal entries in row r are equal to $\frac{\partial}{\partial b_r}G_u + \frac{\partial}{\partial b_t}G_u$, $u \neq r, t$. These are positive given Lemma 3(2). The off-diagonal entries in row w are equal to $\frac{\partial}{\partial b_r}G_u + \frac{\partial}{\partial b_w}G_u$, $u \neq t, w$. Again these are positive except when $u = r$. (c) The sum of the column elements are negative for every column. Consider for example column u. The sum of its elements are $\sum_{x\neq r,t,w} \frac{\partial}{\partial b_x}G_u + (\frac{\partial}{\partial b_r}G_u + \frac{\partial}{\partial b_t}G_u) + (\frac{\partial}{\partial b_r}G_u + \frac{\partial}{\partial b_w}G_u)$, in other words, the change in the probability of winning u when all bids are marginally increased (with the constraint that $b_s = b_t + b_w$). Clearly, this cannot increase the probability that the bidder wins bundle s.

We can now proceed as before by operating on columns $u \neq r, w$ until the matrix is reduced to a 2 by 2 matrix (corresponding to the columns r and w). We will then argue that this 2 by 2 matrix has negative entries except at position $(2, 1)$ so its determinant is positive. The claim then follows from the fact that this number must be multiplied by the product of $|K| - 2$ negative numbers.

Formally, define the operator P^u ($u \neq r, w$) on $\mathbf{M}$ that adds α_{xu} time row u onto row x, $x \neq u$, such $M_{xu} + \alpha_{xu}M_{\mathrm{uu}} = 0$ (note $\alpha_{xu} \geq 0$ and $\sum \alpha_{xu} \leq 1$). The resulting matrix has a zero column at position u except for M_{uu}. Thus $\det \mathbf{M}$ equals M_{uu} times the determinant of matrix $\mathbf{M}$ from which row u and column u have been removed. It is easy to check that this resulting matrix satisfies conditions (a) through (c) above. The only non obvious condition is condition (b). This follows because both $\sum_{x\neq r,t,w} \frac{\partial}{\partial b_x}G_r + (\frac{\partial}{\partial b_r}G_r + \frac{\partial}{\partial b_t}G_r)$ and $\sum_{x\neq r,t,w} \frac{\partial}{\partial b_x}G_r + (\frac{\partial}{\partial b_r}G_r + \frac{\partial}{\partial b_w}G_r)$ are negative (these expressions correspond to the change in the probability of winning r when all bids but w - resp. t - are raised with the constraint that $b_r = b_t + b_w$). Applying operator P^u for all $u \neq r, t, w$ yields the 2 by 2 matrix mentioned above. Hence $\mathrm{sign}(\det \mathbf{M}) = (-1^{|K|-3})$.

Part (1) of Theorem 3 then follows from steps 1 and 2 and the fact that $\mathrm{sign}(\det \boldsymbol{\nabla}\widehat{\mathbf{G}}) = \mathrm{sign}(-1^{|K|})$.

Proof of Part (2): The only non zero multiplier is $\lambda_{\{t\cup w\}=t\cup w}$. From part (1), the costs identified by setting it equal to zero correspond to upper bounds to c_t and c_w and a lower bound to $c_{\{t\cup w\}}$. This identifies the maximum level of synergy consistent with the observed behavior: $\bar{c}_t + \bar{c}_w - \underline{c}_{\{t\cup w\}}$. QED.

9 Appendix B: Data sources and coding issues

9.1 Data sources:

London Buses' tendering program: For each auction and route in the auction, this document provides the tender issue date, the tender return date, the planned start of the contract, the contract duration, together with the start and end point of the routes in the auction.

Bid evaluation documents: These are London Transport Buses internal documents assessing the bids received for one to several routes in a auction. These documents provide information on all route characteristics, including the identity of the incumbent when this is an existing route, the bids received (including combination bids), the identity of the bidders and, most of the time, the garage from which they plan to operate the route.[47] These documents analyze the bids received and make an award recommendation. When this recommendation deviates from the lowest price criterion, the criterion used is detailed and justified.

Route history: History of all the London Bus routes since 1934, compiled by the London Omnibus Traction Society (LOTS). For each route, this data contains information on the identity of the bus bidder, the garage from which operation is carried out, the bus type and peak vehicle requirements (PVR) for weekdays, Saturdays and Sundays. For our analysis we have used weekdays PVR.

Garage history: Document compiled by the London Omnibus Traction Society (LOTS) since the deregulation in 1985. Provides information on openings, closings and transfers of bus garages used for London bus routes. This document is also our primary source of information for entry, mergers and acquisition (secondary sources included London Buses internal memos, companies' websites and LOTS' London Bus and Tram fleetbook publications).

Price deflator: Bids are deflated (Dec. 1995) based on the LTB inflation index formula for the bus services, a weighted index for the retail petroleum prices (5%), labor prices (65%) and retail price index (30%). The data comes from the UK Energy Trends and the UK Monthly Digest of Statistics.

[47] Missing values for the garage locations were completed using the bidder's closest garage to any of the end points of the route at the time of the tender return.

9.2 Coding issues:

Route Alternatives: London Transport Buses sometimes specifies alternative specifications for a route (different bus types, frequencies or routing, for example). By convention, we have coded only the bid information related to the awarded service specification.

Age of vehicle: Vehicle age is the only dimension of the offer, besides price, that is not specified by London Buses. Hence, bidders often submit different bid - vehicle age combinations. In the data, we have coded the bids for both existing and new buses. However, we did not find evidence that would suggest a trade-off between age and bid levels in the award decision. Rather, London Transport Buses seems to evaluate bids holding the age dimension constant, and award decisions are in practice indistinguishable from the award decisions of a contracting authority that would randomize between the age category it prefers, and then selects the best bid within that category. As a result, strategic interactions between the bids along the age dimension can be ignored, and in our empirical analysis, we have focused on the bids submitted for the age category that has attracted bids from the greatest number of bidders.

Auctions: By definition, an auction is a set of routes auctioned at the same time. For our analysis, we have split several of the original auctions into independent subauctions when the following criteria were satisfied: (1) The two subsets of routes were in distinct geographical areas of London, (2) No combination bids were submitted across the two subsets of routes, *and* (3) The bids received on the two different subsets of routes originate from two different sets of bidders, or at least from two different sets of garages.

9.3 Computational Details

We describe here the practical details of the computational approach.

Winning probabilities: As described in section 4, the winning probabilities can be constructed in two ways: (i) simulated and (ii) simulated and smoothed. Simulated winning probabilities result in a step function, and with sufficiently many simulation draws the approximation will be accurate. Simulated and smoothed winning probabilities are additionally treated with a kernel smoother. The simulated and smoothed winning probabilities are suitable for optimization algorithms as they can be evaluated quickly and yield a smooth objective function.

We use the simulated winning probabilities with a large number of simulation draws of $50,000$ in the cost inference. In the construction of the objective function in Figures 1 and

2, and in the validity check of the first order condition, we use simulated and smoothed winning probabilities based on $5,000$ simulation draws.

The computational details of the construction of winning probabilities is the following: We draw opponents' bids from the estimated normal bid density using the normal random variable generator built into the software package GAUSS. We then determine the cheapest allocation for all possible subsets of routes from the set of low opponents' bids following equation (1). We repeat this process $5,000$ times for simulated and smoothed winning probabilities and $50,000$ times for simulated winning probabilities. These simulations give us the pseudo data to assess winning chances.

The simulated winning probabilities are then the empirical cumulative distribution function based on the simulated pseudo data. The function is a step function.

To avoid discrete jumps in the winning probabilities, the simulated and smoothed winning probabilities use the simulated pseudo data with a kernel estimator. The kernel estimator uses a product of normal kernels and a bandwidth equal to the standard deviation of low opponents bid prices for all items times $2/25$. The bandwidth is chosen based on visual inspection.

Optimization: In the consistency check of the first order conditions and in the construction of Figure 2, we calculate optimal bids for each observation by using the optimization algorithm supplied in the software package GAUSS. In the consistency check of the first order condition, we start the algorithm at three distinct and randomly chosen values. In the construction of the figure, we experiment with alternative starting values based on adjacent points to find the maxima.

Cost inference: We employ simulated winning probabilities based on a sample of $50,000$ simulated data points. To calculate the gradient of the probability of winning, we take the differential evaluated at a one percent change in the bid vector. If the resulting derivative is zero, we take the differential evaluated at a five percent of the bid. The inverse matrix of the gradient is calculated numerically using the software package GAUSS.

References

[1] Armstrong, M. (1996), Multi-product Nonlinear Pricing, Econometrica, 64(1), 51-75.

[2] Armstrong, M. and J.C. Rochet (1999), Multi-dimensional Screening: A User's Guide, European Economic Review, 43, 959-979.

[3] Armstrong, M. (2000), Optimal Multi-Object Auctions, Review of Economic Studies, 67, 455-481.

[4] Athey, S. and P. Haile (2002), Identification of Standard Auction Models, Econometrica, 70(6), 2107-2140.

[5] Athey, S and P. Haile (2004), Nonparametric Approaches to Auctions, forthcoming in the Handbook of Econometrics, Volume 6.

[6] Athey, S., J. Levin and E. Seira (2004), Comparing Open and Sealed Bid Auctions: Theory and Evidence from Timber Auctions, mimeo.

[7] Ausubel, L., P. Cramton, R.P. McAfee and J. McMillan (1997), Synergies in Wireless Telephony: Evidence from the Broadband PCS Auctions, Journal of Economics and Management Strategy, 6(3), 497-527.

[8] Avery, C. and T. Hendershott (2000), Bundling and Optimal Auctions of Multiple Products, Review of Economic Studies, 67, 483-497.

[9] Berman, Simeon (1963), Note on Extreme Values, Competing Risks and Semi-Markov Processes, Annals of Mathematical Statistics, 34(3), 1104-1106.

[10] Billingsley, P. (1995), Probability and Measure, John wiley and Sons, 3rd ed.

[11] Branco, F. (1997), Multi-object Auctions with Synergies, mimeo.

[12] Campo, S, I. Perrigne and Q. Vuong (2003), Asymmetry in First-Price Auctions with Affiliated Private Values, Journal of Applied Econometrics, 18, 197-207.

[13] Clarke, E. (1971), Multipart pricing of public goods, Public Choice, 8, 17-33.

[14] Cramton, P, Y Shoham and R. Steinberg (2006), Combinatorial Auctions, Cambridge University Press.

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

[16] Donald, S., H. Paarsch and J. Robert (2005), An Empirical Model of the Multi-Unit, Sequential, Clock Auction, forthcoming Journal of Applied Econometrics.

[17] Elyakime, B., J.J. Laffont, P. Loisel and Q. Vuong (1994), First-Price Sealed-Bid Auctions with Secret Reservation Prices, Annales d'Economie et de Statistique, 34, 115-141.

[18] Epstein, R, L. Henriquez, J. Catalan, G. Weintraub and C. Martinez (2002), A Combinatorial Auction Improves School Meals in Chile, Interfaces, 32(6), 1-14.

[19] Fevrier, P., R. Preget, and M. Visser (2004), Econometrics of Share Auctions, technical report, INRA, Paris.

[20] Griffiths D. (1980), Interval Estimation for the Three-Parameter Lognormal Distribution via the Likelihood Function, Applied Statistics, 29, 58-69.

[21] Groves, T. (1973), Incentives in Teams, Econometrica, 41, 617-631.

[22] Guerre, I., I. Perrigne and Q. Vuong (2000), Optimal Nonparametric Estimation of First-Price Auctions, Econometrica, 68(3), 525-574.

[23] Haile, P. and E. Tamer (2003), Inference from an Incomplete Model of English Auctions, Journal of Political Economy, 111(1), 1-51.

[24] Hansen, L. (1982), Large Sample Properties of Generalized Method of Moments Estimators, 50, 1029-1053.

[25] Hortacsu, A. (2002), Mechanism Choice and Strategic Bidding in Divisible Good Auctions: An Empirical Analysis of the Turkish Treasury Auction Market, University of Chicago, mimeo.

[26] Hortacsu, A. and S. Puller (2005), Testing Models of Strategic Bidding in Auctions: A Case Study of the Texas Electricity Spot Market, technical report, Department of Economics, University of Chicago.

[27] Jackson, M., L. Simon, J. Swinkels and W. Zame (2002), Communication and Equilibrium in Discontinuous Games of Incomplete Information, Econometrica, 70(5), 1711-1740.

[28] Jofre-Bonet M. and M. Pesendorfer (2003), Estimation of a Dynamic Auction Game, Econometrica, 71(5), 1443-1489.

[29] Judd, K. (1998), Numerical Methods in Economics, MIT Press: Cambridge.

[30] Kastl, J. (2005), Discrete Bids and Empirical Inference in Divisible Good Auctions, technical report, Department of Economics, Northwestern University.

[31] Krasnokutskaya, E. (2004), Identification and Estimation in Highway Procurement Auctions under Unobserved Auction Heterogeneity, technical report, University of Pennsylvania.

[32] Krishna, V. and R. Rosenthal (1996), Simultaneous Auctions with Synergies, Games and Economic Behavior, 17(1), 1-31.

[33] Krishna, V., and M. Perry (2000), Efficient Mechanism Design, mimeo available at http://econ.la.psu.edu/~vkrishna/papers/vcg20.pdf

[34] Levin, J. (1997), An Optimal Auction for Complements, Games and Economic Behavior, 18, 176-192.

[35] Laffont, J.J., H. Ossard and Q. Vuong (1995), Econometrics of First-Price Auctions, Econometrica, 63(4), 953-980.

[36] Laffont, JJ and Q. Vuong (1996), Structural Analysis of Auction Data, American Economic Review, Papers and Proceedings, 36, 414-420.

[37] Li, Tong, Isabelle Perrigne and Quang Vuong (2002), Structural Estimation of the Affiliated Private Value Auction Model, Rand Journal of Economics, forthcoming.

[38] Manski, C and E. Tamer (2002), Inference on Regressions with Interval Data on a Regressor or Outcome, Econometrica, 70, 395-409.

[39] Maskin, E and J. Riley (2000), Equilibrium in Sealed High Bid Auctions, Review of Economic Studies, 67, 413-438.

[40] McAdams, David (2005), Identification and Testable Restrictions in Private Multi-Unit Auctions, mimeo

[41] McAfee, R.P., J. McMillan and M.D. Whinston (1989), Multiproduct Monopoly, Commodity Bundling and Correlation of Values, Quaterly Journal of Economics, 102(2), 371-383.

[42] McFadden, D. (1989), A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration, Econometrica, 57(5), 995-1026.

[43] Milgrom, P. (2000), Putting Auction Theory to Work: The Simultaneous Ascending Auction, Journal of Political Economy, 108(2), 245-272.

[44] Moreton, P. and P. Spiller (1998), What's in the Air: Interlicence Synergies and their Impact on FCC Broadband PCS License Auctions, Journal of Law and Economics.

[45] Pakes, A. and D. Pollard (1989), Simulation and the Asymptotics of Optimization Estimators, 57(5), 1027-1057.

[46] Palfrey, T. (1983), Bundling Decisions by a Multi-product Monopolist with Incomplete Information, Econometrica, 51(2), 463-483.

[47] Schmalensee, R. (1984), Gaussian Demand and Commodity Bundling, Journal of Business, 57(1), S211-S230.

[48] Silverman, B.W. (1986), Density Estimation for Statistics and Data Analysis, Chapman and Hall.

[49] Smith, R.L. (1985), Maximum Likelihood Estimation in a Class of Nonregular Cases, Biometrika, 72, 67-90.

[50] Vickrey, W. (1961), Counterspeculation, Auctions, and Competitive Sealed Tenders, Journal of Finance, 16, 8-37.

[51] Whinston, M.D. (1989), Tying, Foreclosure, and Exclusion, American Economic Review, 80(4), 837-859.

[52] Wolak, F. (1993) Identification and Estimation of Cost Functions Using Observed Bid Data, An Application to Electricity Markets, in: Dewatripont, Hansen and Turnosky, Proceedings of the Eigth World Congress of the Econometric Society, Cambridge University Press.