Combination Bidding in Multi-Unit Auctions*

Estelle Cantillon[†]and Martin Pesendorfer[‡]

July 17, 2002

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, because of anticipated synergies, bidders are allowed to submit bids on combinations of routes as well as on individual routes. We show that equilibrium combination bidding 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 identification in the multi-unit first price auction. In particular, we show that the presence of combination bids is a necessary condition for identification. We propose an estimation method to infer the multidimensional private information. The method consists of two stages. In the first stage, the distribution of bids is estimated parametrically. In the second stage, costs and the distribution of costs 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 evaluate the welfare impacts of allowing combination bids in that market.

Preliminary and incomplete. Comments welcome.

[†]Harvard Business School and CEPR. Email: ecantillon@hbs.edu.
[‡]London School Economics, CEPR and NBER. Email: martin.pesendorfer@yale.edu.

1

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 - where bidders submit bids contingent on the final allocation - allow bidders to transmit richer information regarding the value they attach to the objects for sale. When the objects are not independent, for instance, because the bidders value bundles of objects differently than the sum of the constituent parts, allowing such contingent bids is a necessary condition for efficiency and optimality (Groves, 1973 and Clarke, 1971; Levin, 1997, Armstrong, 2000 and Avery and Hendershott, 2000).

This was well understood by the London bus procurement authorities. Indeed, two of the motivations for allowing combination bids in the London bus routes market were that they would allow bidders to pass on, through lower bids, some of the cost savings resulting from cost synergies between routes, and at the same time, that they would enhance the efficiency in the allocation of routes across bidders.

However, allowing combination bids in the first price auction may also hurt efficiency and costs. In section 2, we introduce a model of a private values multi-unit procurement auction that allows for cost synergies between objects. We highlight two distinct motivations for combination bidding. On the one hand, 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 synergies. The reason is that combination bids allow bidders to link otherwise independent markets and leverage any advantage they may have in one market into the other. This effect may increase procurement costs and hurt efficiency. On the other hand, 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 whether the leverage effect or the synergy effect dominates.

Section 3 provides conditions under which the unobserved costs and cost distribution are identified from bid data. We assume that bidders assess their winning probabilities correctly and choose bids to maximize their profits. We show that a necessary condition

2

for identification of the multidimensional private information based on bid data is that the auction permits bidders to submit a full set of combination bids, in addition to stand-alone bids. Moreover, as long as bidders do make use of all their bids, the cost distribution function is identified. The intuition is that identification requires the set of allowable bids – which represents the observed information – to be of the same dimension as the private information to infer. In particular, this means that the general first price multi-unit auction model with only stand-alone bids is not identified. Finally, we show that any constraint on bids, such as reserve prices, or the condition that combination bids must be lower than the sum of the constituent stand-alone bids, can reduce the dimensionality of the observed information and therefore introduce a level of underidentification. A cost range rather than a cost point is identified. We characterize this cost range.

Section 4 proposes our estimation method. It is based on Guerre, Perrigne and Vuong (2000)'s two stage estimation method for single unit auctions but extends it to multi-unit auction environments. The estimation proceeds in two stages. In the first stage, the multivariate joint distribution of bids for all units is estimated parametrically. In the second stage, the multivariate cost distribution is inferred using the first order conditions for optimal bids.

Section 5 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 relative importance of the leverage and synergy motivations for combination bids. Our very preliminary results are reported in section 6. We calculate the percentage mark-up of bids (relative to cost) for a selected sample of contracts for which stand-alone and combination bids are submitted. We find that the mark-up is about 26% on stand-alone bids and 35% on combination bids. Calculating the synergy effects we find that the cost of a combined route is on average 8% lower than the sum of the costs for the individual routes. However, we also find evidence that not all combination bids are motivated by underlying cost synergies.

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. 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.

3

Exceptions include the sequential auctions analyzed by Donald, Paarsch and Robert (2001) and Jofre-Bonet and Pesendorfer (2001) and the discriminatory multi-unit auction analyzed by Hortacsu (2002). Donald et al. and Jofre-Bonet and Pesendorfer's approaches generalize previous estimation techniques to account for intertemporal linkages between auctions. Hortacsu (2002) studies the Turkish Treasury bill auctions. He shows that bidders' valuation schedules are identified from their observed demand schedules in the discriminatory multi-unit auction and proposes an estimation strategy based on resampling techniques.

There has also been a number of recent theoretical analyses of multi-unit auctions. 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 first price auction. Our leverage motivation 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.

2 The Model

This section introduces the model and highlights its key properties. The model integrates the salient features of the London bus routes market.

A seller simultaneously offers m contracts for sale to N risk neutral bidders. Each bidder i privately observes a cost draw, $c_s^i \in \mathbb{R}$, for each possible subset of the contracts, $s \subseteq S = \{1, \ldots, m\}$. Notice that there are a total of $2^m - 1$ possible subsets of contracts. We say that contracts s and t, with $s \cap t = \emptyset$, are *independent* from bidder i's perspective if $c_s^i + c_t^i = c_{s \cup t}^i$ where $c_{s \cup t}^i$ denotes bidder i's cost for the combination of contracts s and t.

They are *complements* if $c^i_{s \cup t} < c^i_s + c^i_t$ and *substitutes* if $c^i_{s \cup t} > c^i_s + c^i_t$.

Contract costs are ex-ante distributed according to the joint distribution $F((c^i_s)_{s \subseteq S, i=1,...N} | X)$ where $X = (x, w)$ denotes a vector of observable contract characteristics x and bidder characteristics w. We assume that F is common knowledge, and that it has a bounded and coordinate-wise convex support with a well defined strictly positive density everywhere. Notice that our formulation permits correlation in bidders' costs across bidders and contracts.

We compare two auction rules. The first auction rule replicates the rule used in the London bus routes market. Bidders may submit bids on all subsets of the set of contracts. Let b^i_s denote bidder i's bid on the subset of contracts $s \subseteq S$, and let $b^i = (b^i_1, ..., b^i_s, ... b^i_S) \in \mathbb{R}^{2^m - 1}$. We sometimes use the symbol b^i_{-s} to denote the vector of bids by bidder i on all contracts except for s. Bidders pay the value of their winning bids and the auctioneer selects the winner(s) based on the allocation that minimizes her total payment. Formally, the last restriction requires that $b^i_{s \cup t} \leq b^i_s + b^i_t$ for all s, t such that $s \cap t = \emptyset$. A combination bid must be no greater than the sum of its constituent stand-alone bids. Otherwise the auctioneer would select $b^i_s + b^i_t$ and ignore the combination bid $b^i_{s \cup t}$.[1]

The second auction rule is the standard simultaneous first-price auction where bidders are allowed to submit bids on the individual contracts only. That is $b^i = (b^i_1, ..., b^i_m) \in \mathbb{R}^m$.

Fix bidder i, and for each contract s, define B^{-i}_s as the lowest bid submitted by bidder i's opponents on route (combination) s. By convention, let $B^{-i}_\emptyset = 0$ and let $B^{-i} = (B^{-i}_1, \dots, B^{-i}_S)$.

When only stand-alone bids are allowed, bidder i's bid on an individual contract $s \in S$ competes only against the best bid of his opponents on that particular contract, B^{-i}_s. Nevertheless, his optimization problem departs from the optimization of bidders in a single unit auction because his costs depend on the final allocation of contracts. As a result, his bids on the individual contracts must take into account the possibility that he may also win other contracts.[2]

When combination bids are allowed, a different trade-off arises from the fact that bidders' own bids compete with one another. Formally, with m contracts, there are 2^m possible

---

[1] *Existence of an equilibrium* in the multi-unit auction with combination bidding is guaranteed. There always exists an equilibrium where all bidders submit a bid for the bundle S only. To see this, notice that given that i's opponents only submit a bid for the bundle, any bid by i on contract $s \subset S$ can only win together with i's own bid on $S \backslash s$. In other words, a bidder can only win a contract if he wins all of them. As a result, setting $b^i_s = \infty$ for all $s \neq S$ is a best response. Of course, other equilibria may also exist.

[2] This is the standard "exposure problem."

winning allocations. Either $b_s^i + B_{S\backslash s}^{-i}$ beats any other alternative bid combination $b_t^i + B_{S\backslash t}^{-i}$ for $t \neq s$ and B_S^{-i}, in which case bidder i wins exactly the set of contracts s with cost c_s^i, or $B_S^{-i} < b_s^i + B_{S\backslash s}^{-i}$ for all s, in which case bidder i does not win anything. This yields the following payoff function for bidder i (ignoring ties):

$$\begin{cases} b_s^i - c_s^i & \text{if } b_s^i + B_{S\backslash s}^{-i} < \min\{b_t^i + B_t^{-i} \text{ for } t \neq s, B_S^{-i}\} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

Consider the set of contracts s. Holding the distribution of the opponents' best bids $(B_1^{-i}, ..., B_S^{-i})$ fixed, decreasing b_s^i increases bidder i's chance to win exactly set of contracts s by lowering the price of allocation $b_s^i + B_{S\backslash s}^{-i}$ relative to the others. However, this may come at the expense of winning potentially more profitable contract combinations if for instance $b_s^i - c_s^i < b_t^i - c_t^i$ for some t.

This is a standard price discrimination trade-off and it is analogous to the pricing problem of the multi-product monopolist in the multi-dimensional screening literature. To make this analogy more transparent, suppose that there are only two contracts, 1 and 2. Figure 1 represents bidder i's bid $(b_1^i, b_2^i, b_{1\cup 2}^i)$ in the (B_1^{-i}, B_2^{-i}) space (in that space, combination bid $b_{1\cup 2}^i$ can be represented by a line with slope -1).
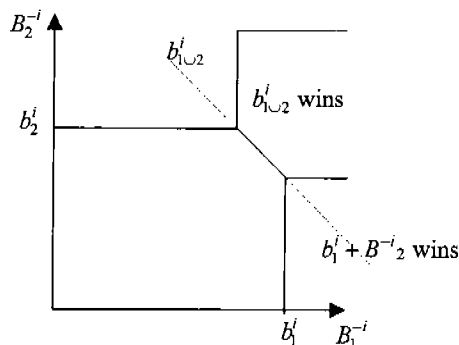
Figure 1:

Ignoring $B_{1\cup 2}^{-i}$, bidder i wins contract 1 only when $b_1^i + B_2^i < \min\{b_{1\cup 2}^i, B_1^{-i} + b_2^i\}$, that is whenever the realization of (B_1^{-i}, B_2^{-i}) falls in the lower right quadrant of figure 1. Similarly, bidder i wins contract 2 only when (B_1^{-i}, B_2^{-i}) is in the upper left quadrant. In the upper right truncated quadrant (shaded), $b_{1\cup 2}^i$ beats every other bid combination so bidder i wins both goods. He wins none in the lower left quadrant.

Replacing the distribution of the opponents' best bids (B_1^{-i}, B_2^{-i}) by the distribution of consumers reserve prices for good 1 and good 2, and reinterpreting b_1^i, b_2^i and $b_{1\cup 2}^i$ as

the prices for goods 1, 2 and the bundle of the two respectively, yields the standard multiproduct pricing problem. McAfee, McMillan and Whinston (1989) use figure 1 to derive a *sufficient* condition for bundling to be profitable for the monopolist with additive costs. In particular, they find that when demand is independent across goods – in our setting, whenever (B_1^{-i}, B_2^{-i}) are independently distributed – submitting a price for the bundle that is lower than the sum of the individual prices is optimal for the monopolist. Armstrong and Rochet (1999) solve for the global maximization of the multi-product monopolist. Their analysis confirms that bundling (the equivalent of submitting a non trivial combination bid $b_{1\cup2} < b_1 + b_2$) is profitable unless there is strong correlation across buyers' reservation values. See also Armstrong (1996 and 2000).

The following example suggests that the strategic effect to which combination bidding gives rise may have perverse consequences on welfare.

Example: The Leverage Motivation

Consider two independent contracts. Types are one-dimensional and independently distributed. Let $c_s^i(\theta^i)$ be bidder i's cost for contract s, for the realization of type $\theta^i \in \mathbb{R}$. Without loss of generality, we take θ^i to be *i.i.d* uniform on [0,1] since any asymmetry among bidders can be captured by the c_s^i functions. There are 3 bidders: A, B and C. Bidders A and B are "local" bidders in the sense that bidder A is only interested in contract 1, $c_1^A(\theta) < \infty$ but $c_2^A(\theta) = \infty$ for all θ and bidder B is only interested in contract 2. By contrast, bidder C is interested in both contracts and he is a "global" bidder. Furthermore, assume that $c_1^A(\theta) = c_1^C(\theta) = c_1(\theta)$ and $c_2^B(\theta) = c_2^C(\theta) = c_2(\theta)$ for all θ,[3] costs are increasing in θ, $c_1'(.)$ and $c_2'(.) > 0$, and that there is no synergy between the two contracts for the global bidder, $c_{1\cup2}^C = c_1^C + c_2^C$.

It is instructive to first consider the scenario where combination bidding is not allowed. Since the allocation in each market is independent of the outcome and the bids in the other market, and given that bidders are symmetric at the individual market level, the unique equilibrium is symmetric and in strictly increasing strategies. As a result, the outcome is efficient. Moreover, conditional on the optimal reserve price and the usual regularity condition, this simple auction format minimizes procurement costs.

Now suppose that bidders are allowed to submit combination bids, and towards a contradiction suppose there is an equilibrium where bidders only submit stand-alone bids. From the perspective of the global bidder, $(b_1^A, b_2^B) = (B_1^{-C}, B_2^{-C})$ are independently distributed. Therefore, the analysis of McAfee et al. (1989) applies and bidder C will find it advanta-

---

[3]This means that bidders are symmetric at each object level.

7

geous to submit a non trivial combination bid at equilibrium, $b_{1\cup2}^C < b_1^C + b_2^C$. This means that we can rule out the equilibrium where bidders only submit stand-alone bids (with the "trivial" combination bid $b_{1\cup2}^C = b_1^C + b_2^C$). *Combination bidding must take place in any equilibrium.*

In this example, combination bidding hurts efficiency and cost. Efficiency is hurt because whether the local bidder A wins contract 1 or not, no longer depends on bidder A's and bidder C's signals only but also on bidder B's signal (through the combination bid of the global bidder). By the revenue equivalence theorem, and given that the optimal auction is efficient, procurement cost is also higher.[4] The example can be generalized and we have the following result.

Proposition 1 *Suppose that the contracts are independent and that competition in the local markets is symmetric. Then allowing combination bids can only increase procurement costs and hurt efficiency.*

Proposition 1 suggests a class of environments where allowing combination bidding hurts both costs and efficiency. Nevertheless, we have also constructed examples, especially when high synergy levels are present, in which equilibrium bidding with combination bids results in both higher efficiency *and* lower procurement costs relative to the equilibrium outcome of the game in which combination bidding is not allowed.

There are three lessons from this analysis. First, observing a combination bid lower than the sum of bids for the stand-alone constituent units is no guarantee that there are underlying synergies. Submitting a combination bid can be profitable exactly for the same reason why the multi-product monopolist finds price discrimination profitable.

Second, correlation in the environment is an important determinant of combination bidding. In fact, we can show that there is no independent role for combination bidding if bidders' private information is unidimensional (that is, if costs are perfectly correlated across contracts) and bidders are not too asymmetric.

Third, understanding the costs and benefits of combination bidding is an important policy question. One benefit of combination bidding is that it allows bidders to better

---

[4]The intuition is that the combination bid pools the two markets together and it allows the global bidder to leverage any advantage he has in one market into the other. Indeed, suppose that bidder A has a very high cost realization for contract 1 that is, the global bidder has an advantage in market 1. Then, if the global bidder only submits a combination bid, it reduces the toughness of the competition he faces in the second market because bidder B needs to have a really low bid to compensate bidder A's high bid and have a chance to win. This mechanism, market linkage through combination bidding, is analogous to the leverage theory in industrial organization (Whinston, 1989).

align bids on costs. In fact, combination bidding is a necessary condition for efficiency and optimality (Groves, 1973 and Clarke, 1971; Levin, 1997, Armstrong, 2000 and Avery and Hendershott, 2000). However, our analysis shows that combination bidding may also have perverse effects in the first price auction. This suggests that the question is ultimately an empirical one because the answer depends on the nature and extent of synergies present in the market. This motivates the next section.

3 Identification

This section describes our identification results for the multi-unit first price auction. We observe data on bids, contract characteristics and bidder characteristics. Our goal is to infer costs, which we do under two assumptions: (1) the observed data on bids, contract characteristics and bidder characteristics can be used to correctly infer bidders' beliefs about the winning chances of their bids, and (2) bidders choose bids to maximize the interim expected payoff.

The bidding model is identified if the distribution of costs can be uniquely inferred from the observed data. In this section, we provide a new positive result for non-parametric identification in the private values multi-unit first price auction and show that the model is generically identified when combination bids are allowed. We also illustrate how to obtain identification bounds for the cost parameters when additional restrictions are placed on the set of allowable bids.

Guerre, Perrigne and Vuong (2000) prove non-parametric identification in the one dimensional independent private values setting for single object first-price auctions.[5] Hortacsu (2002) studies the homogeneous multi-unit discriminatory auction for Treasury bills. His identification problem is closer to ours since it is intrinsically multi-dimensional: he observes (a distribution of) demand schedules and his goal is to infer the (distribution of) marginal valuation curves. Extending Guerre et al., Hortacsu proves non-parametric identification in the case where bidders submit demand functions. The difference between our setting with heterogeneous goods and Hortacsu's model of homogeneous and divisible goods is that demand is identified by a *vector* of costs $(c_1, ..., c_S) \in R^{2^m-1}$ in our setting whereas it is identified by a marginal valuation *function* in his. This leads to different mathematical structures.

---

[5]Other results for single unit auctions include Laffont and Vuong (1996) and Li, Perrigne and Vuong (2000) who extend Guerre et al.'s identification result to affiliated private values, and Athey and Haile (2001) who analyze the identification problem when some bid observations are missing.

9

3.1 Identification Conditions

We start with the general (unconstrained) combinatorial first price auction. Bidders submit bids on all subsets of the set of objects, the auctioneer selects the cheapest bidder-bid allocation and the winners pay the price of their winning bids. Let $h(b^1, ..., b^N | X)$ denote the equilibrium joint distribution of bids and let $\mathcal{B}$ be its support.[6] We make the following assumption on h and $\mathcal{B}$:[7]

Assumption 1: $(b^1, ..., b^N) \in R^{2^m - 1 \times N}$ is distributed continuously on a closed and full dimensional support.

Fix bidder i and let B^{-i} be the $2^m - 1$ dimensional vector of best bids by bidder i's opponents on each bundle. Given m objects, there are 2^m possible winning allocations of the objects between bidder i and his opponents. Either $b_s^i + B_{S \backslash s}^{-i}$ for some $s \subseteq S$ is the cheapest bidder-bid combination, in which case bidder i wins bundle s, or B_S^{-i} is, in which case bidder i does not win anything. Let $G_s(b^i | \beta^i, X): R^{2^m - 1} \to [0, 1]$ denote the (correctly inferred) probability that bid vector b^i by bidder i wins exactly bundle s conditional on his actual submitted (equilibrium) bid being β^i and on some covariates X.[8]

In the multi-unit auction, bidders submit bids on all subsets of objects and therefore solve a $2^m - 1$ dimensional problem (in b_s^i, $s \subseteq S$):

$$\max_{(b_s^i)_{s \subseteq S}} \sum_{s \subseteq S} (b_s^i - c_s) G_s(b^i) \tag{2}$$

(for simplicity, we omit from now on the conditioning variables β^i and X in the expression of G_s). A direct consequence of assumption 1 (proved in the appendix) is that G_s is continuous and a.e. differentiable in b_t^i for all t. Hence the expression in (2) is continuous and a.e. differentiable. At any point where it is differentiable, the optimal bid vector by bidder i

---

[6]The joint distribution of bids is trivially identified from the data when all bids are observed. When not all bids are observed, additional identification conditions are required. See Athey and Haile, 2001 for the single object case.

[7]In principle, the assumption of equilibrium behavior imposes some restrictions on the observed distribution of bids. In the absence of a full characterization of equilibrium behavior in the combinatorial first price auction, assumption 1 should be seen as an additional assumption made on the observed data.

[8]This expression for the probability of winning recognizes that the distribution of bids across bidders may be correlated. Though empirical papers on auctions often assume independence across bidders, identification results based on the first order condition do not require such assumption. For a recent example of extension of identification in the single object first price auction to affiliated values, see Li, Perrigne and Vuong, 2000.

10

must satisfy the first order conditions:

$$G_t(b^i) + \sum_{s \subseteq S}(b_s^i - c_s)G_s^t(b^i) = 0 \qquad t \subseteq S$$

or, in matrix notation:

$$\nabla G(b^i)[b^i - c] = -G(b^i) \tag{3}$$

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

The first order conditions define a *system of linear equations* in the unknown cost parameters, c_s, $s \subseteq S$. Identification of a cost realization then reduces to the question of existence and uniqueness of a solution to this system.

Proposition 2 (Sufficient condition for identification) *Suppose that assumption 1 holds and define the $(2^m - 1)$ by $(2^m - 1)$ matrix $\nabla G(b^i)$ by $\nabla G_{t,s}(b^i) = G_s^t(b^i)$. A sufficient condition for identification in the first price multi-unit auction is that $\nabla G(b^i)$ is invertible for all i and all b^i.*

Proof. Since the first order conditions in (3) define a system of linear equations in the unknown cost parameters (the $[b^i - c]$ vector), the invertibility of matrix $\nabla G(b^i)$ is a necessary and sufficient condition for a unique solution $c^i = \phi^i(b^i) \in \mathbb{R}^{2^m - 1}$. Identification of the distribution of costs F follows directly. Indeed, if the system in (3) admits a unique solution $\phi^i(b^i)$ for all b^i and all i, then the distribution of bids $h(b^1, ..., b^N)$ defines a unique distribution of costs $F(c^1, ..., c^N) = \int_{\{(b^1, ..., b^N)|\phi^i(b^i) \leq c^i \text{ for all } i\}} h(b)db$. ∎

To provide sufficient conditions for the matrix ∇G to be invertible, the following definition will be useful:

Definition (irrelevant bids): Bid b_s^i by bidder i is *irrelevant* if $G_s(b_s^i, b_{-s}^i) = 0$ and there exists $\varepsilon > 0$ such that $G_s(b_s^i - \varepsilon, b_{-s}^i) = 0$. Otherwise it is *relevant*.

Irrelevant bids never win. Nevertheless, irrelevant bids can be optimal from a bidder's perspective because of the leverage motivation for combination bids: submitting a bid that never wins on a contract ensures that this bid does not compete with any other, potentially more profitable, bid.[9] Such bids are problematic for inference. Indeed, suppose bidder i submitted an irrelevant bid on contract s. Then, any alternative bid vector $(\widehat{b_s^i}, b_{-s}^i)$ with

---

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

$\widehat{b}_s^i > b_s^i$ would have been equally optimal for bidder i, and therefore equally informative. More formally, the definition of irrelevant bids implies that $G_t^s(b^i) = 0$ (small changes in b_s^i do not affect the probability that bidder i wins contract s or any other contracts) and $G_s^t(b^i) = 0$ for all t. This means that the row and column corresponding to contract s in matrix ∇G are all zeros. Therefore ∇G cannot be inverted.

Nevertheless, we can still use the information that winning contract s was not profitable for bidder i.

Definition (effective bid): Fix the distribution of $(b^1, ..., b^N)$. Consider any b_s^i for some bidder i and bundle s. Its associated effective bid, b_s^{eff} is the highest bid on contract s such that b_s^{eff} is relevant subject to the condition that $b_s^{eff} \leq b_s^i$.

If b_s^i is irrelevant, b_s^{eff} is such that $G_s(b_s^{eff}, b_{-s}^i) = 0$ but $G_s(b_s^{eff} - \varepsilon, b_{-s}^i) > 0$ for all $\varepsilon > 0$. Since payoffs are continuous in bids, bidder i is indifferent between submitting (b_s^i, b_{-s}^i) or (b_s^{eff}, b_{-s}^i). In addition, since bidder i found it profitable not to win contract s, it must be that:

$$\sum_{t \neq s}(b_t^i - c_t)G_t^s(b_s^{eff}, b_{-s}^i) + (b_s^{eff} - c_s)G_s^s(b_s^{eff}, b_{-s}^i) = \mu_s \geq 0 \tag{4}$$

where all derivatives are left derivatives.[10] We shall see in the next section how this inequality can be used to provide bounds on the unobserved parameters. In the meantime, we prove the following properties of the elements of matrix ∇G:

Lemma 1 (Properties of ∇G) *Suppose that assumption 1 holds. Consider matrix ∇G (with elements defined by $\nabla G_{t,s} = G_s^t$) evaluated at any optimal bid vector b^i by bidder i. Then:*

(1) $G_t^t \leq 0$ for all t, and strictly so if b_t^i is strictly relevant.

(2) $G_s^t \geq 0$ for all $t \neq s$.

(3) $\sum_s G_s^t \leq 0$ for all t, and strictly so for some t as long as b^i includes at least one strictly relevant bid.

Proposition 3 (Necessary and sufficient conditions for the invertibility of ∇G) *Suppose that assumption 1 holds. Then ∇G is invertible at any optimal bid vector b^i if and only if all bids in b^i are relevant. Moreover, 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.*

---

[10]We consider the left derivatives because the critical value b_s^{eff} corresponds to a potential point of non-differentiability in the payoff function (see remark 1 following proposition 5 in the appendix).

Lemma 1 and proposition 3 are proved in the appendix. Proposition 3 has a number of direct implications for identification in multi-unit auctions.

Corollary 1 *(i) Suppose that assumption 1 holds and that all equilibrium bids are strictly relevant, then the private values combinatorial first price auction model is identified. (ii) If only stand-alone bids are permitted, then the model is not identified. (iii) If contracts are independent and only stand-alone bids are permitted, then the model is identified.*

Corollary 1 is proved in the appendix. A general lesson is that identification requires the dimensionality of the observed information to match that of the information to be inferred. Specifically, in the multi-unit auction model with multi-dimensional types, the underlying private information to infer (the costs c_s) is $2^m - 1$ dimensional. On the other hand, the set of bids determines the dimensionality of the *observed* information. When bidders make full use of all their bids, as in corollary 1(i), the observed information is $2^m - 1$ dimensional, and identification follows. By contrast, when only stand-alone bids are permitted, the observed information is m dimensional and there is no hope to infer costs, unless the dimensionality of private information is also m.[11]

3.2 Incorporating Restrictions of the London Bus Market

The previous section suggested that the presence of irrelevant bids may be problematic for identification. In addition, real-life auctions include various restrictions on the set of allowable bids which may also lead to a reduction in the dimensionality of the observed information and therefore the violation of the conditions in corollary 1(i).

In this section, we show how to extend our identification results and derive identification bounds in these cases. We illustrate our approach by considering three types of restrictions present in our data. First, the rule of the auction imposes that bids on a combination of contracts must be no greater than the sum of the constituent bids. Second, London Transport Buses imposes a (secret) reserve price. Third, bidders are not obliged to submit bids on all routes and some bidders indeed submit bids only on a subset of the routes auctioned in any particular tranche. Our interpretation is that it was not profitable for these bidders to submit a bid that would have had a positive chance of winning.

---

[11]In addition, if the inference ignores the multi-unit nature of the auction (by treating each object as a separate auction), it can be shown that the cost estimates are biased downward if contracts are complements and upward if they are substitutes.

Together, these restrictions mean that the optimization problem that any typical bidder faces becomes:

$$\max_{b^i} \sum_{s \subseteq S} (b_s^i - c_s) G_s(b^i) \tag{5}$$

subject to a combination bid constraint:

$$b_s^i \leq b_w^i + b_t^i \text{ for all } s, t, w \subseteq S \text{ such that } t \cap w = \emptyset \text{ and } t \cup w = s \tag{6}$$

and a reserve price constraint *in the event bidder i does submit a bid on any given contract*:

$$b_s^i \leq R_s \text{ for all } s \subseteq S \tag{7}$$

The fact that bidders do not need to submit a bid on all contracts introduces a discontinuity in the payoff function. Moreover, because of the multi-unit nature of the auction, we cannot transform problem (5) subject to (6) and (7) into an optimization problem of a *continuous function* over a *compact set* (for example, by imposing compulsory bidding) such that any solution to the original problem is a solution to the transformed problem.[12]

Nevertheless, we can still use the logic of the standard approach to optimization under constraints and use first order conditions to infer costs. We proceed as follows. First, we view reserve prices on each subset of the routes as "bids" submitted by the auctioneer. As a result, though $(b^1, ..., b^N)$ may no longer be viewed *stricto sensu* as distributed continuously (given the mass point generated by non submitted bids), the probabilities of winning remain continuous and a.e. differentiable.[13] So does bidder i's optimization problem. In addition, with reserve prices viewed as bids, the definition of effective bids extends straighforwardly to the presence of reserve prices.

Second, whenever a bidder submits a relevant bid on all contracts, the Kuhn-Tucker conditions for the optimization problem (5) subject to (6), (7) *and* compulsory bidding provide a proper description of the bidder's optimization problem.

---

[12]To see this, suppose that at the optimum b^* of (5), bidder i submits a relevant bid on contract s but not on contract t. Then, his first order condition with respect to b_s^i takes into account the fact that b_s^i does not have any externality on contract t: $G_t^s(b^*) = 0$. However, this property does not necessarily hold if all bids satisfy the reserve price constraint. As a result, b_s^* may no longer satisfy the first order condition in the transformed problem.

[13]Alternatively, assumption 1 should be interpreted as a restriction on *submitted* bids.

Third, whenever one of these bids, say b_s^i, is irrelevant (including because it was not submitted), the first order condition with respect to that particular contract can be evaluated at (b_s^{eff}, b_{-s}^i) where $b_s^{eff} < b_s^i$ is the associated effective bid to b_s^i that makes bidder i's bid on bundle s just relevant.

To summarize, given any observed bid vector b^i, we propose to evaluate bidder i's first order condition with respect to contract s at (b_s^{eff}, b_{-s}^i) where b_s^{eff} is the effective bid associated with b_s^i. Clearly, effective and actual bids only differ in the cases of non submitted bids or irrelevant bids. These are also the cases where, as we have argued, the first order condition needs to be evaluated at the value of the bid that makes it just relevant. In addition, since $G_s^t(b_s^i, b_{-s}^i) = 0$ for all $t \neq s$, and for all s such that b_s^i is irrelevant (or non submitted), these adjusted first order conditions can still be expressed as a system of linear equations of the form

$$\nabla G(b^i)[b^{eff} - c] = D \tag{8}$$

where the ∇G matrix is now defined as

$$\nabla G_{t,s}(b^i) = G_s^t(b_s^{eff}, b_{-s}^i),$$

the $[b^{eff} - c]$ vector is now evaluated at the associated effective bid vector, and the D column vector collects the G functions with the Lagrangian multipliers:

$$D_s(b^i; \mu, \lambda) = -G_s(b_s^{eff}, b_{-s}^i) - \mu_s - \sum_{r,t} \lambda_{s=r \cup t} + \sum_{r,t} \lambda_{t=s \cup r}$$

μ_s is positive if bidder i did not submit a bid or submitted an irrelevant bid on contract s, and $\lambda_{s=r \cup t} \geq 0$ is the Lagrangian multiplier for the constraint $b_s^i \leq b_r^i + b_t^i$ for $s = r \cup t$ and $r \cap t = \emptyset$.

The same arguments as in proposition 3 can be used to prove that matrix ∇G with rows evaluated at (b_s^{eff}, b_{-s}^i) is invertible. This means that

$$[b^{eff} - c] = \nabla G^{-1}(b^i) D(b^i; \mu, \lambda) \tag{9}$$

is the solution to (8). Expression (9) is important. It says that, given any fixed values for the Lagrangian multipliers, costs are uniquely identified from the bid observation. In particular, if no constraint is binding, all the Lagrangian multipliers are equal to zero and costs are identified: we are back to the special case of corollary 1. By contrast, any binding constraint introduces a degree of underidentification because we do not observe the value of the Lagrangian multiplier that solved bidder i's optimization problem when he (optimally)

15

chose to submit bid vector b^i. The only thing we know from the previous discussion is that the multiplier of the binding constraint is positive. The next proposition characterizes more precisely the extent of the underidentification in the cost parameters:

Proposition 4 (Identification bounds) *Any binding constraint introduces a one-dimensional degree of underidentification in the cost vector c with the following properties: (1) c_s depends positively and linearly on the value of the multipliers μ_s; but it is independent of μ_t for all $t \neq s$; (2) c_s depends positively on $\lambda_{s=t\cup w}$, and negatively on $\lambda_{t=s\cup w}$ for all s, t and w.*

The proof of proposition 4 can be found in the appendix. It uses Cramer's rule and the properties of determinants to sign how the solution to (9) varies with the value of the Lagrangian multipliers.

There are several important elements to note in proposition 4. First, irrelevant bids and non submitted bids only affect the identification of the cost parameter of the associated contract. 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 or non submitted bids, bids on *other* contracts do not affect the probability that bidder i wins the contract on which he either did not submit a bid, or 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 between the first order condition with respect to the contract on which submitting an irrelvant bid is optimal and the other first order conditions. A direct consequence is that, in the event of non submitted bids, we can infer that the costs of the contract was higher that the reserve price – exactly as in the single unit first price auction.

Second, proposition 4 allows us to derive bounds on the value of the cost parameters that can rationalize the observed bids. For example, suppose that constraint $b^i_s \leq b^i_t + b^i_w$ is binding. The solution to (9) evaluated at $\lambda_{s=t\cup w} = 0$ provides a lower bound to the cost parameter c_s. Of course, given the analysis in section 2, a parameter of greater interest still is the extent of synergies between contracts t and w. The next result is proved in the appendix.

Corollary 2 (The maximum level of synergies is identified) *Consider any 2 disjoint contracts, t and w. If the combination bid constraint for these contracts is binding at the optimum, an upper bound to the synergy involved between these two contracts is given by*

16

the solution $c_t + c_w - c_{t\cup w}$ of the system in (9) when the Lagrangian multiplier $\lambda_{\{t\cup w\}=t\cup w}$ is set equal to zero.

3.3 Identification of the Cost Distribution Function [to be added]

4 Estimation Method

This section describes our estimation approach. Section 4.1 describes our parametric density specification. Section 4.2 describes a simulated method of moments estimator for the bid density function. Section 4.3 takes the bid density function as given and describes our numerical method to infer costs.

We observe data on a cross section of auctions $t = 1, \dots T$. Let b^{it} denote the bid vector of bidder i submitted for the contracts in auction t and let $X^t = (x^t, w^{1t}, \dots, w^{Nt})$ denote the route and bidder characteristics on auction t. $w^{-i,t}$ denotes the vector of characteristics for bidders other than bidder i and we sometimes write $X^{it} = (x^t, w^{it}, w^{-i,t})$ where the i superscript indicates that bidder characteristics are evaluated from bidder i's perspective. In this section, we make the following further assumption on the data generation process.

Assumption. The cost realizations for bidder i are stochastically independent of the cost realizations of bidder j, for all $j \neq i$, conditional on characteristics X^t.

Furthermore, we assume that characteristics X^t are observable to the bidders and the econometrician. We do not consider bidder or contract heterogeneity that is not observed to the econometrician.

4.1 A Multivariate Bid Density Function

We specify the density function of (latent) bids b^* of bidder i in auction t as a multivariate log-normal density function in which the parameters are a (linear) function of bidder and auction characteristics X^{it}. Given our independence assumption on the cost draws across bidders conditional on X^{it}, the bids by bidder i are stochastically independent from the bids of bidder j conditional on characteristics X^{it}.

The statistical model for latent bids by bidder i on auction t is:

$$
\begin{bmatrix}
\ln(\frac{b_1^*}{IC_1} - \gamma) \\
\vdots \\
\ln(\frac{b_S^*}{IC_S} - \gamma)
\end{bmatrix}
= \mu(X^{it}) + \Lambda(X^{it}) \cdot \varepsilon
\tag{10}
$$

17

In (10), latent bids are normalized by an internal cost estimate IC_s, μ denotes the $(2^m - 1)$ dimensional vector of means, Λ is a $(2^m - 1)$ by $(2^m - 1)$ matrix and ε is a $(2^m - 1)$ dimensional standard normal random variable: $\varepsilon \sim N(0, I)$ with 0 denotes the null vector and I denotes the identity matrix. We assume that the mean μ_s and Λ_{st} are linear functions of characteristics,

$$\mu_s(X^i) = \beta_s X^{it} \quad \text{and} \quad \Lambda_{s\tau}(X^{it}) = \alpha_{s\tau} X^{it}$$

Thus, $\beta = [\beta_1, \dots, \beta_S]$ is a $(2^m - 1) \times k$ dimensional parameter matrix where k denotes the number of explanatory variables in X and $\alpha = [\alpha_{s\tau}]$ is a $(2^m - 1) \times (2^m - 1)$ dimensional matrix. Each element $\alpha_{s\tau}$ is a parameter vector of dimension k. Notice, that our specification implies that the variance-covariance matrix of log bids is given by $\Sigma = \Lambda \Lambda'$.

Latent bids above the reserve price are not observed. As in Laffont, Ossard and Vuong (1995), we define the observed bid as equal to the reserve price, R_s^t, when the latent bid is not observed:

$$b_s^{it} = b(b^* | X^{it}) = b_s^{*it} \cdot 1_{\{b_s^{*it} < R_s^t\}} + R_s^t \cdot 1_{\{b_s^{*it} \geq R_s^t\}}$$

Henceforth, we use this convention and restrict attention to observed bids.

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 method of moments. Assuming that the lower bound γ is known, the regularity conditions of maximum likelihood are satisfied, and maximum likelihood yields consistent and efficient estimates of the parameters (α, β) as the number of auctions T gets large. If the lower bound γ is not known and is to be estimated, then a regularity condition of maximum likelihood is violated. The method of moments provides an alternative estimation method that yields consistent estimates for the parameter vector $\theta = (\alpha, \beta, \gamma)$, as the number of auctions T increases (see Hansen (1982)). Numerical calculation of the moment conditions can be computationally intensive as our density is multivariate. Simulation estimators (McFadden (1989) and Pakes and Pollard (1989)) provide an elegant solution to this problem. The next section describes our estimator.

4.2 A Simulated Method of Moments Estimator

Consider the difference between the observed and the (conditional) expected kth moment of the bid vector:

18

$$v^{itk} = (b^{it})^k - E\left[b^k | X^{it}, \theta\right]$$

The symbol $(b^{it})^1$ denotes the vector of first moments $(b_1^{it}, \ldots, b_S^{it})$, the symbol $(b^{it})^2$ denotes the vector of second moments $((b_1^{it} b_1^{it}), b_1^{it} b_2^{it}, \ldots, (b_S^{it} b_S^{it}))$ and so on. Notice, that the difference v^{itk} when evaluated at the true parameter value θ^*, is mean independent of the exogenous data:

$$E\left[v^{itk} | X^{it}, \theta = \theta^*\right] = 0$$

Given this condition together with some standard regularity conditions, we can adopt the method of moments estimator described in Hansen (1982). The estimating equation is given by

$$(b^{it})^k = E\left[b^k | X^{it}, \theta\right] + v^{itk}$$

Unfortunately, $E\left[b^k | X^{it}, \theta\right]$ 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. The expected kth order moment of the observed bid can be written as:

$$E\left[b^k | X^{it}, \theta\right] = \int \cdots \int \left[(b(b^*(\varepsilon | X^{it}, \theta) | X^{it}, \theta)^k \frac{\phi(\varepsilon)}{\varphi(\varepsilon)}\right] \varphi(\varepsilon) d\varepsilon_1 \cdots d\varepsilon_S$$

by multiplying and dividing the integrand by the importance function $\varphi(\cdot)$ and where ϕ denotes the multivariate standard normal density for $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_S)$, and the function b^* is implicitly defined in equation (11).[14]

Given a fixed set of random draws, ε_{it}, for each bidder and auction, we can calculate an estimate, $\widehat{b^k}$, with the property:

---

[14]

$$b^*(\varepsilon_1, \ldots, \varepsilon_S | X^{it}) = IC \cdot \left[\exp\left(\mu(X^{it}) + \Lambda(X^{it}) \cdot \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_S \end{pmatrix}\right) + \begin{pmatrix} \gamma \\ \vdots \\ \gamma \end{pmatrix}\right]$$

19

$$E_\varepsilon \widehat{b^k}(X^{it}, \theta, \varepsilon) = E\left[b^k | X^{it}, \theta\right]$$

The estimate leads to a new estimating equation

$$b^{itk} = \widehat{b^k}(X^{it}, \theta, \widehat{\varepsilon}) + \widehat{v}^{itk} \tag{11}$$

Since $\widehat{b^k}$ is an unbiased estimator of $E\left[b^k | X^{it}, \theta\right]$, the new prediction error, $\widehat{v}^{itk}$, is also mean independent of the exogenous data at the true parameter values. This suggests a method of moments technique may still be appropriate to estimate θ. Under regularity conditions which are satisfied here, McFadden (1989) and Pakes and Pollard (1989) show that this suggestion is correct.

Our estimator of the above form is found by taking, for each bidder and auction, independent draws from the multivariate importance function φ of errors $\varepsilon = (\varepsilon^{11}, \dots, \varepsilon^{NT})$. As importance function φ we use the standard multivariate normal denisty ϕ. A simulated moment estimator is defined as the average across multiple simulated draws to reduce the variance of the estimate while preserving unbiasedness. This is, we take L draws, $\varepsilon = (\varepsilon^1, \dots, \varepsilon^L)$ and calculate

$$\widehat{b^k}(X^{it}, \theta, \widehat{\varepsilon}) = \frac{1}{L} \sum_{l=1}^{L} \left[(b(b^*(\varepsilon_1^{itl}, \dots, \varepsilon_S^{itl} | X^{it}, \theta) | X^{it}, \theta)^k \frac{\phi(\varepsilon_1^{itl}, \dots, \varepsilon_S^{itl})}{\varphi(\varepsilon_1^{itl}, \dots, \varepsilon_S^{itl})} \right]$$

This equation for $\widehat{b^k}$ can then be substituted into the estimating equation (12). Let $\widehat{v}^{it}$ denote the resulting vector moment prediction errors for bidder i on auction t. It consists of the $(2^m - 1)$ vector of bids for individual route combinations stacked on top of the $(2^m - 1)(2^m)/2$ elements of cross products of bids for individual route combinations, and so on. We denote a typical element of the vector by $\widehat{v}^{itl}$.

Let W^{itk} be a matrix of instruments for the kth moment prediction error $(k = 1, \dots, K)$. We can write the instrument matrix for bidder i on auction t, W^{it}, as a block diagonal matrix:

$$W^{it} = \begin{bmatrix} W^{it1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & W^{itK} \end{bmatrix}$$

20

The method of moment estimation technique is based on the restriction that $\hat{v}^{it}$ is uncorrelated with the exogenous data W^{it}. With NT independent observations, the sample analog of this restriction involves the sample correlation

$$\bar{g}(\theta) = \sum_i \sum_t \sum_l g_{itl}(\theta) = \sum_i \sum_t \sum_l \hat{v}^{itl}(\theta) \otimes W^{itl}$$

where the vector of instruments W^{itl} has at least as many elements as there are parameters in θ. An estimate $\hat{\theta}$ is chosen to minimize a quadratic distance measure

$$G(\theta) = \bar{g}(\theta)' A \bar{g}(\theta)$$

for some positive definite matrix A. A preliminary estimate $\hat{\theta}_1$ is obtained by setting A equal to the identity matrix. Then, a new weighting matrix is calculated as the sample variance of the individual moment conditions, $g_{itl}(\hat{\theta}_1)$. A second and final estimate, $\hat{\theta}$, is then obtained from the use of this moment condition. The variance of $\hat{\theta}$ is estimated by the formula given, for example in 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 θ.

4.3 Inference of the Cost Distribution Function

This section describes our numerical approximation technique to infer the cost distribution function based on estimates of the bid density.

Let $h(b_1^i, ..., b_s^i, ..., b_S^i | X^i)$ denote the estimated probability density function of bids by bidder i on all the subsets of S conditional on X^i. The probability that bidder i's bid (vector) b^i wins exactly route (combination) s conditional on X^i, $G_s(b^i | X^i)$, can be written as a function of the density of the lowest bids by i's opponents conditional on X^i, which we denote as $h_{(1)}(B_1^{-i}, ..., B_S^{-i} | X^i)$. Notice further that the density $h_{(1)}(B_1^{-i}, ..., B_S^{-i} | X^i)$ can be expressed directly as a function of the bid density $h(.|.)$. 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 $h(.|X^j)$ conditional on the characteristics X^j. We then determine the low competitors'

bids $(B^{-i}|X^i) = (B_1^{-i}, ..., B_S^{-i}|X^i)$. We repeat this exercise L times by repeatedly drawing bids and determining the low competitors' bids. The "pseudo data" of competitors' low bids, $(B_l^{-i}|X^i)_{l=1}^L$, is then used directly to approximate the probability that bidder i wins exactly route s with the bid b^i, $G_s(b^i|X^i)$. The empirical frequency of this event is given by:

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

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

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

Next, we describe how we determine the cost density and distribution function. We approximate the cost density function using a step function on a specified grid. We partition the bid (and cost) space into M intervals, $C_s = \{[\frac{(i-1)R_s}{M-1}, \frac{iR_s}{M-1})\}_{i=1}^{M-1} \cup [R_s, \infty)$.[15] The first $M-1$ intervals are of length $\frac{R_s}{M-1}$ each. The last interval, $[R_s, \infty)$, accounts for unobserved bids. We take the Cartesian product of the partitions in each dimension and specify the grid in the bid space as $C = X_{s \in S} C_s$. Observe that the grid C consists of $n = M^{2^m-1}$ cubes and each cube is of dimension $2^m - 1$. The fineness of the grid is determined by the number M, and our numerical approximation becomes more accurate as we increase M. We denote a typical element of C by C_l. The probability of the cube C_l in the bid space is denoted by h_l. It is given by:

$$h_l(X^i) = \int_{C_l} h(b|X^i) db.$$

In order to determine the associated probability of the cube C_l in the cost space, we conduct the following calculation: For each $l = 1 \ldots n$, we select a bid vector $b_l \in C_l$. Without loss of generality we chose b_l as the midpoint of the cube C_l. The previous section describes how to obtain the cost range associated with b_l based on the Kuhn-Tucker conditions for optimal bids. We denote the associated cost range with $\phi_l(b_l|X^i)$. Notice, that

---

[15]For simplicity of exposition, we assume that the cost space equals the bid space. This assumption is not required and can be relaxed.

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 ∞. The value of the cost density function on the cube C_l equals the frequency with which cost observations fall into the cube. If the cost range is a singleton, then the frequency is exactly determined, while if the cost range is not a singleton, then this translates into an upper and a lower bound on the cost distribution function only.

We specify the density function of the cube C_l for a lower and an upper bound of the cost distribution function using the componentwise maximum and minimum of the cost range,

$$f_l^{\text{inf}} = \frac{1}{n} \sum_{k=1}^{n} 1_{\{\sup(\phi_k(b_k|X^i)) \subset C_l\}} \cdot h_k(X^i),$$

$$f_l^{\text{sup}} = \frac{1}{n} \sum_{k=1}^{n} 1_{\{\inf(\phi_k(b_k|X^i)) \subset C_l\}} \cdot h_k(X^i).$$

where $\sup(\phi)$ (and $\inf(\phi)$) are the componentwise largest (and smallest) element in the set ϕ.

Our estimator of the cost distribution function is then the cumulative distribution function associated with the lower and upper bound on the empirical frequency densities. They are given by:

$$F^{\text{inf}}(c_1, ..., c_S|X^i) = \sum_{l=1}^{n} f_l^{\text{inf}} \cdot 1_{\{C_l \subset \{c' | c_1' \leq c_1, ..., c_S' \leq c_S\}\}},$$

$$F^{\text{sup}}(c_1, ..., c_S|X^i) = \sum_{l=1}^{n} f_l^{\text{sup}} \cdot 1_{\{C_l \subset \{c' | c_1' \leq c_1, ..., c_S' \leq c_S\}\}}.$$

We can make three observations: First, our numerical approximation entails an error and the error becomes negligible as M and L increase. Second, we calculate the standard errors for cost estimates using the delta method. Third, if no constraint is binding, then the upper and lower bound on the cost distribution function coincide.

5 The London Bus Market

This section describes the London bus market, gives descriptive summaries of our data and provides evidence for our chosen specification.

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 $900 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 (the set of routes that corresponds to an invitation is called a *tranche*). 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 operators may submit sealed bids for individual routes. In addition, operators may submit a bid for route combinations within the tranche. A bid specifies an annual price at which the operator 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

---

[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.

24

follows EU law for fair competition.[18] In practice, this means that the contract is awarded to the low bidder but deviations at the margin are possible to account for operator quality for instance.[19] To allow winning operators to reorganize and order new buses if necessary, contracts start 8 to 10 months after the award date.

Description of the bid data.[20] We have collected data on 179 tranches consisting of a total of 674 routes offered to operators between December 1995 and May 2001 (return date). For each tranche and for each route in the tranche, 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, mobility route); annual mileage; bus type (single deck, midibuses, double deck or routemaster); and the peak vehicle requirement;[21] (3) the identity of bidders and all their submitted bids (including bids for combinations of the routes in the tranche). For the auctions held starting in May 2000, the data also contain an internal cost estimate generated by London Transport Buses for every route.

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. 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 internal cost estimate. The internal cost estimate explains 90% of the variation in the bids and 92% in the variation of low 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 tranches in our data that have no more than 3 routes. Table 1 provides summary statistics of our bid data

---

[18]EEC directive 93/38.

[19]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.

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

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

[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.

for these tranches.[23]

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

Variable	Obs	Mean	Std	Min	Max
PVR	218	9.65	8.26	1	52
ln(mileage)	218	12.00	1.79	6.21	14.40
ln(ICE)	218	13.28	1.29	10.82	15.56
No-Bidders-per-Tranche	118	3.70	1.74	1	8
No-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

On average 3.7 bidders submit at least one bid on a tranche. The number of bidders ranges between 1 and 8. Fewer bids are submitted on individual routes. On average 2.94 operators 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 a tranche with three routes or less. Of those, 26 win a contract.

Operators submit a total of 641 stand-alone bids. The distribution of 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 tranches 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 operators to offer discounts for combinations of routes include the possibility to share spare vehicles and depot 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

---

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

[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.

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. Bus operators tend to be organized in groups with operational subsidiaries active in local areas of London. As of November 2000, there were 64 pre-qualified bus companies in the London area. Because operational companies that belong to the same group don't bid against one another, we define the bidding entity at the group level, and refer to it as a "bidder" or an "operator." This yields 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 bus operator active in the tendered bus services in London, we have a complete history of its depots (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, by bus type). Depots are leased on a long term basis or bought, and a typical garage has capacity for 50-100 buses and serves about 8 routes. Table 2 provides descriptive statistics of our operator data for the period between November 1995 and May 2001.

Table 2 Descriptive Statistics for the Operators (Monthly, Nov 95 - May 01)

Variable	Obs	Mean	Std	Min	Max
Bidders-with-active-garage	67	20.06	1.11	18	23
Active-garages	67	82.75	3.47	76	88
Garages-per-operator	1,344	4.12	5.22	1	25
Buses-per-operator	1,344	267.90	376.77	0	1,355

A few elements are worth noting. First, asymmetry among operators is considerable. For example, in November 2000, a total of 10 operators had one garage, 4 operators had two garages, one operator had 6 garages, one operator has 7 garages, one operator had 9 garages, one operator had 11 garages, one operator have 13 garages, and one operator has 21 garages. This size asymmetry is also reflected in the distribution of market shares in our sample as well as in the range of bus types bidders operate.

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 tranche, submitted 10.6% of all the bids, and bidders with only one established garage submitted another 15.95% of the bids. In our

27

sample, there was an entrant or a bidder with only one garage bidding on 49.11% of the routes. The incumbency rate is 62%. Such active fringe would make collusion very difficult to sustain.

The determinants of uncertainty and bidder participation. A measure of uncertainty is the relative difference between the lowest and second lowest stand-alone bid, or money left on the table. It equals 13.54%. Thus, stand-alone bids overpay by about 110,500 Pounds on average. This suggests considerable uncertainty about the competitors' bids. 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. Thus, even as the number of bids submitted increases, the uncertainty does not vanish. With five or more bidders the amount overpaid for the average contract equals almost 63,000 Pounds.

What determines the uncertainty in bids? At the operator 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 operators concerning the expected cost of labor or fuel (there are well functioning markets for these), but opportunity costs may not be known to other operators. 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, tranche, or route level. In other words, does the opportunity cost vary at the firm level, the garage level or route level? To examine these questions we decompose the variation in the bid submission decision.

We examine how much of the variation in the decision variable is explained by tranche fixed effects, route fixed effects, tranche-depot and tranche-operator 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 tranches 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 3 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.[25] For example, the test statistic for the hypothesis that tranche fixed effects are zero (model (2)) is an F- distributed random variable with (67, 3288) degrees of freedom.

Table 3 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+Tranche Fixed Effects (TF)	70	0.28	0.27	3.00**
(2')	DM+Route Fixed Effects	170	0.30	0.26	1.63**
(3)	DM+TF+ Operator Fixed Effects	92	0.31	0.29	4.14**
(4)	DM+TF+Depot Fixed Effects	159	0.40	0.37	6.83**

* Tranches 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 tranche or at the route level. Both do equally well on the basis of the adjusted $\overline{R}^2$. We also tested whether route fixed effects are significantly different from zero when tranche 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 tranche fixed effects are accounted for*. 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 tranche fixed effects. We may interpret this uncertainty as a bidder specific idiosyncracy arising at the route and tranche 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 tranche fixed effects. Second, a substantial part of the uncertainty in bidders' decisions is explained by bidder asymmetry captured by dead mileage, bidder fixed effects and depot

---

[25] ** denotes we can reject the null hypothesis at a 1% confidence level.

fixed effects. Third, there is considerable residual uncertainty for each bidder arising at the route and tranche 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 tranche is the proper level of analysis for this market. First, the temporal simultaneity of the auction for the routes in the same tranche, their geographic proximity as well as the existence of combination bids requires that we analyse them at the same time. Second, several elements suggest that inter-tranche 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 tranches tend to cover different geographical areas.[26] *First price*: The competitive assumption is consistent with the existence of the active fringe of bidders. *Private Values:* Most of the inputs used by operators 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 does allow for cost correlation across routes, it does not assume it. This seems important in view of the evidence presented in table 3.

6   Estimation Results [preliminary and incomplete]

This section describes our estimates.

Our data do not contain reserve prices. To account for the reserve price in the estimation, we presume that reserve prices follow a specific functional form that is linear in the internal cost estimate. A lower bound on the reserve price is then the highest ratio of accepted bid to the reserve price. In our data this ratio equals 1.45. We use this number for our analysis. We vary this number to assess the robustness of our estimates to changes in the reserve price rule. Specifically, lowering the number should not affect the estimates qualitatively.

---

[26]The geographic dispersion of the tranches together with local nature of the business reduces the interactions among tranches and bidders. We calculated that an average bidder in our sample bid on a tranche every 5 months.

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 exceed the number of parameters. For the mean of route combination s on auction t for bidder i we select the following eight instruments: internal cost estimate, the number of routes in the tranche, dead mileage of bidder i to route (combination) s, the number of garages of bidder i, a dummy that equals one if the route combination s consists of routes with identical bus types[27], the number of firms with a garage within 5 miles of one route on the tranche, a constant, the square of the internal cost estimate. For the second moment of bids of bidder i on route combination $s \cup t$ we select the internal cost estimate for the route combination $s \cup t$, the total number of routes within the tranche, dead mileage of bidder i to route combination $s \cup t$, a dummy that equals one if the route combination $s \cup t$ consists of routes with identical bus types and a constant. The total number of instruments exceeds the number of parameters in the model which guarantees identification.

Potential bidders on auction t include all bidders with a garage within 5 miles of at least one route within auction t. Bidders who submit a bid and have a garage further than 5 miles from all routes on auction t are classified as fringe bidders. We assume that it is common knowledge that fringe bidders submit a bid on at least one route on auction t and include fringe bidders in the set of potential bidders.

Parameter Specification: There are a number of natural restrictions to impose on the way the parameters of the bid density depends on bidder and auction characteristics $X^{it} = (x^t, w^{it}, w^{-i,t})$. First, the parameters should be invariant with respect to permutation in the characteristics of bidder i opponents, $w^{-i,t}$. Second, the parameters should also be invariant with respect to permutations of the indices of sets of routes with the same number of routes. This implies that $\beta_s = \beta_t$, for all s, τ such that $|s| = |\tau|$. Restrictions on the matrix Λ are $\alpha_{s,s} = \alpha_{\tau,\tau}$, $\alpha_{s,\tau} = \alpha_{\tau,s}$, $\alpha_{s,u} = \alpha_{\tau,u}$ 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 tranche, bidder, contract heterogeneity and synergy effects. As explanatory variables we include the internal cost estimate[28],

---

[27]The dummy also equals one for stand-alone routes.

[28]As explained above, the data provide information on the internal cost estimates for the most recent 60 routes. For earlier route we construct a measure of the predicted internal cost estimate. The predicted internal cost estimate is calculated based on a regression of the internal cost estimate on route characteristics. Route characteristics included in the regression are the number of peak vehicle required, the mileage, and bus type. The regression explains 95% of the variation in the internal cost estimate.

the number of bidders with at least one garage within 5 mile of one of the routes of the tranche, the number of garages of bidder i within 5 miles of one route on the tranche, the dead mileage of bidder i. This yields the following specification for μ_s^i:

$$\mu_s^i = \beta_1 + \beta_2 \text{ICE}_s + \beta_3 \text{NO-GARAGES-TOTAL} + \beta_3 \text{NO-GARAGES-i}$$
$$+ \beta_4 \text{DEAD-MILEAGE-i}_s + \beta_5 \text{IDENTICAL-BUS-TYPES}_s + \beta_6 \text{NO-ROUTES-IN-THE-TRANCHE}$$

We specify the elements $\alpha_{s,\tau}$ of the matrix Λ as follows:

$$\alpha_{s,\tau} = \lambda_1 1_{\{s=\tau\}} + \lambda_2 1_{\{s\neq\tau \text{ and } s\cap\tau=\emptyset\}} + \lambda_3 1_{\{s\neq\tau \text{ and } s\cap\tau\neq\emptyset\}} + \lambda_4 1_{\{\text{routes } s \text{ and } \tau \text{ have the same bus type}\}}$$

The first constant accounts for diagonal elements in Λ, while the second and third constant account for off diagonal elements. We distinguish two off-diagonal effects depending on whether route s and τ have a non-empty interesection or not. The last constant accounts for synergy effects that may arise if the bus types required in routes s and τ are the same.

7   Conclusions [To be written]

8   Appendix A: Data sources and coding issues

8.1   Data sources:

London Buses' tendering program: For each tranche and route in the tranche, 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 tranche.

Bid evaluation documents: These are London Transport Buses internal documents assessing the bids received for one to several routes in a tranche. 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.[29] These documents analyze the bids received and make an award recommendation. When

---

[29]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.

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 operator, 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.

Depot history: Document compiled by the London Omnibus Traction Society (LOTS) since the deregulation in 1985. Provides information on openings, closings and transfers of bus depots 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).

8.2 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, operators 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 main regressions, we have focused on the bids submitted for the age category that has attracted bids from the greatest number of bidders.

Tranches: By definition, a tranche is a set of routes auctioned at the same time. For our analysis, we have split several of the original tranches into independent subtranches 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,

33

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 Appendix B: Proofs

Proposition 5 (Continuity and differentiability of G_s) *Suppose bids are distributed continuously on a closed support of full dimension. Then bidder i's probability of winning contract s, $G_s(b^i)$, is continuous at all b^i on its domain for all s, and it is a.e. continuously differentiable in b^i_t for all t.*

Proof: Given m contracts, there are 2^m possible winning allocations among bidder i and his opponents: either $b^i_s + B^{-i}_{S\backslash s}$ is the bidder-bid combination most advantageous to the auctioneer, or B^{-i}_S is. Ignoring ties, bidder i wins contract s if

$$b^i_s + B^{-i}_{S\backslash s} < \min\{b^i_t + B^{-i}_{S\backslash t} \text{ for all } t \neq s; B^{-i}_S\} \tag{12}$$

Define $W^i_s(b^i_{-s}) = \min\{b^i_t + B^{-i}_{S\backslash t} \text{ for all } t \neq s; B^{-i}_S\} - B^{-i}_{S\backslash s}$. (notice that W^i_s is a function of b^i_{-s}).

Claim 1: $W^i_s(b^i_{-s})$ is a random variable distributed continuously on a closed support.

Proof: For W^i_s not to be a random variable, it must be $\min\{b^i_t + B^{-i}_{S\backslash t} \text{ for all } t \neq s; B^{-i}_S\}$ is perfectly correlated with $B^{-i}_{S\backslash s}$. Because all the $B^{-i}_{S\backslash t}/B^{-i}_S$ are best bids on different subsets of the contracts, the only way for them to be perfectly correlated on a one by one basis involves perfect correlation of some of the bids in the $(b^1, ..., b^N)$ space. This is ruled out given the full dimensionality of the bids support. Hence W^i_s is a random variable. It has a closed support since $(b^1, ..., b^N)$ have a closed support, and it is continuously distributed on that support since $(b^1, ..., b^N)$ are distributed continuously on its support. QED

Let $\mathcal{W}^i_s$ denote the support of W^i_s. Define $f^i_s(.|b^i_{-s})$ as the probability distribution function of W^i_s on $\mathcal{W}^i_s$ and $f^i_s(.|b^i_{-s}) = 0$ on $R\backslash\mathcal{W}^i_s$. Then

$$G_s(b^i) = \int_{b^i_s} f^i_s(x|b^i_{-s})dx$$

is well defined and it represents the probability that bidder i wins bundle s given his submitted vector of bids b^i. By construction, it is continuous in b^i_s and differentiable at any value of b^i_s where $f^i_s(.|b^i_{-s})$ is continuous. (e.g. Rudin, 1967, thm 6.20). By claim 1 and by construction of f^i_s, the only points of discontinuity of f^i_s in b^i_s lie at the boundaries of the

support of W_s^i. Therefore G_s is a.e. differentiable in b_s^i. Continuity and a.e. differentiability with respect to b_t^i, $t \neq s$, is proved in a similar fashion since the distribution of W_s^i is defined implicitly through integration over $B_{S\backslash s}^{-i} \backslash B_S^{-i} \backslash B_{S\backslash t}^{-i}$ (cfr expression (12)). This implies that $f_s^i(.|b_{-s}^i)$ is differentiable a.e. in b_t^i, $t \neq s$. Therefore, so is G_s^i. QED

Remark 1 *Because f_s is always left and right continuous, left and right derivatives are always well defined.*

Remark 2 *Points of non-differentiability of G_s and irrelevant bids. A noteworthy consequence of the proof of proposition 5 is that G_s fails to be differentiable in b_s at b_s^{eff} that makes the bid on bundle s just relevant (since b_s^{eff} corresponds to the upper bound of W_s^i). At b_s^{eff}, G_s may also fail to be differentiable with respect to a bid on one other bundle.*

Proof of lemma 1: Given m contracts, there are 2^m possible allocations of these contracts between bidder i and his opponents: Either $b_s^i + B_{S\backslash s}^{-i}$ is the best bid combination submitted, in which case bidder i wins exactly subset s of the contracts, or B_S^{-i} is the best bid combination, in which case bidder i does not win anything.

Any increase in b_t^i makes allocation $b_t^i + B_{S\backslash t}^{-i}$ more expensive relative to the other ones, but otherwise does not affect the relative ranking of $b_s^i + B_{S\backslash s}^{-i}$, $s \neq t$, and B_S^{-i}. Hence the probability that any of these competing allocations wins cannot decrease: $G_s^t \geq 0$ for $t \neq s$ and $\sum_s G_s^t \leq 0$ for all t. Likewise, the probability that allocation $b_t^i + B_{S\backslash t}^{-i}$ wins cannot increase: $G_s^s \leq 0$.

If b_t^i is strictly relevant (in particular, this means that $G_t > 0$), G_t^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 profits bidder i makes from his other bids), a contradiction with the fact that b^i is optimal for bidder i.

We now show that $\sum_s G_s^t < 0$ for some t when b^i contains at one least strictly relevant bid. Towards a contradiction, suppose that $\sum_s G_s^t = 0$ for all t. This means that the support of $\min\{b_S^i, b_t^i + B_{S\backslash t}^{-i} \text{ for all } t \neq s\}$ is distinct from the support of B_S^{-i} (by proposition 4, these random variables are distributed continuously on a closed support). Since one of bidder i's bids wins sometimes, it means that[30]

$$\min\{b_S^i, b_t^i + B_{S\backslash t}^{-i} \text{ for all } t \neq s\} \leq B_S^{-i}, \text{ for all realizations of } B^{-i} \tag{13}$$

---

[30]The expression in (13) implicitely assumes that the supports of $\min\{b_S^i, b_t^i + B_{S\backslash t}^{-i} \text{ for all } t \neq s\}$ and B_S^{-i} are convex. It is straightforward to adapt the argument to non convex (and closed) support.

Since b^i is optimal

$$\max_{B^{-i}}\{\min\{b^i_S - B^{-i}_S, \ b^i_t + B^{-i}_{S\setminus t} - B^{-i}_S \text{ for all } t \neq s\}\} = 0$$

(for otherwise increasing all bids by epsilon would be a profitable deviation). But this means that B^{-i}_S competes at least with one of bidder i's bids. Hence $\sum_s G^t_s < 0$ for at least one t. QED

Proof of proposition 3:

From proposition 5, we know that the point at which a bid becomes just relevant, b^{eff}_s, is a potential point of non differentiability of the G_s function. Nevertheless these remain both left and right differentiable and we adopt the convention that whenever a bid is just relevant, its derivatives is taken from the left (i.e. making it strictly relevant). With this convention, lemma 1 also applies to all relevant bids.

From now on, we also introduce the notation "contract 0" with the convention that $b^i_0 = 0$, $S\setminus\{0\} = S$, and say that bidder i wins contract 0 when B^{-i}_S is the winning bidder-bid combination (bidder i wins nothing). With this definition, $G_0 = 1 - \sum_{s\subseteq S} G_s$.

The following definition will be useful:

Definition: The set of contracts $\Omega \subseteq S \cup \{0\}$ forms a connected chain of substitutes if for all s and s' in Ω ($s' \neq 0$), either $G^{s'}_s > 0$ or there exist $w_1, ..., w_n \in \Omega$ such that $G^{w_1}_s > 0, G^{w_2}_{w_1} > 0, ..., G^{s'}_{w_n} > 0$.

Claim 1: If all bids on the set of contracts S are relevant, then $\Omega = S \cup \{0\}$ forms a connected chain of substitutes.

Proof: First, since $G^s_s < 0$ (lemma 1(1)), any contract $s \subseteq S$ must be connected with at least one other contract. In addition, contract 0 is connected to at least one other contract by lemma 1(3). Second, 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 now prove that if all bids in S are relevant, 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 continuous random variables distributed on a closed support (using assumption 1 and a similar argument to that of lemma 1), $\min\{b^i_t + B^{-i}_{S\setminus t}, b^i_s + B^{-i}_{S\setminus s}\}$ and $\min_{w\neq t,s}\{b^i_w + B^{-i}_{S\setminus w}\}$. Since all bids are relevant, sometimes $\min\{b^i_t + B^{-i}_{S\setminus t}, b^i_s + B^{-i}_{S\setminus s}\} \leq \min_{w\neq t,s}\{b^i_w + B^{-i}_{S\setminus w}\}$ (bidder i wins contract s or t) and sometimes $\min\{b^i_t + B^{-i}_{S\setminus t}, b^i_s + B^{-i}_{S\setminus s}\} \geq \min_{w\neq t,s}\{b^i_w + B^{-i}_{S\setminus w}\}$. Hence $\min\{b^i_t + B^{-i}_{S\setminus t}, b^i_s + B^{-i}_{S\setminus s}\} = \min_{w\neq t,s}\{b^i_w + B^{-i}_{S\setminus w}\}$ must happen given that their support is closed and that 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. G_w^t or $G_w^s > 0$. A contradiction.

We can now prove that $\det \nabla G < 0$ (so that ∇G is invertible). The proof is by induction. By lemma 1, property (3) holds strictly for at least one contract. We relabel the rows and columns of matrix ∇G such that the sum of the elements in the first row is strictly negative (this does not change the value of the determinant):

$$\sum_s G_s^1 < 0 \tag{14}$$

Consider the linear transformation, L_1 on the columns of ∇G that adds to column $s \neq 1$, α_{1s} times column 1 such that $G_s^1 + \alpha_{1s} G_1^1 = 0$ for $s \neq 1$ (notice, $\alpha_{1s} \geq 0$ and $\sum \alpha_{1s} < 1$ given (14)). This leaves the first row of matrix ∇G with all zeros except in the first position. Denote the resulting matrix by $L_1 \nabla G$ and let $[L_1 \nabla G]$ be matrix $L_1 \nabla G$ from which the first row and the first column have been removed. Since determinants are invariant to linear transformations, $\det \nabla G = \det L_1 \nabla G = G_1^1 \det[L_1 \nabla G]$.

We claim that the resulting $2^m \times 2^m$ matrix $[L_1 \nabla G]$ satisfies properties (1) to (3) of the original matrix, including the strict inequalities. Property (1): The diagonal elements of matrix $[L_1 \nabla G]$ are equal to $G_s^s + \alpha_{1s} G_1^s$. Since the G_t^s elements satisfy properties (1) to (3) and $\alpha_{1i} < 1$, we have $G_s^s + \alpha_{1s} G_1^s < 0$. Property (2): The off-diagonal of the new matrix are equal to $G_t^s + \alpha_{1t} G_1^s \geq 0$ since it is a sum of positive elements. Property (3): The sum of the row elements of the $[L_1 \nabla G]$ matrix is equal to $\sum_{s \neq 1} G_s^t + G_1^t \sum_{s \neq 1} \alpha_{1s} \leq 0$ since $\sum G_s^t \leq 0$ and $\sum_{s \neq 1} \alpha_{1s} < 1$. To show that this inequality holds strictly for at least one row of the new matrix $[L_1 \nabla G]$, we need to consider two cases. First, if any of the elements G_1^s 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 all the elements $G_1^s = 0$ for $s \neq 1$, 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 connected to 0. This means that $\sum_s G_s^t < 0$ in the original matrix, and in the new matrix.

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

To prove the last part of the claim we show that any submatrix made from ∇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. QED

Proof of corollary 1: (i) Follows directly from propositions 2 and 3. (ii) When only stand-

alone bids are permitted, bidders solve the following constrained maximization problem:

$$\max_{b^i} \sum_{s \subseteq S} (b^i_s - c_s) G_s(b) \quad \text{s.t.} \quad b^i_s = b^i_{s \setminus t} + b^i_t \text{ for all } s \subseteq S \text{ and } t \subset s$$

Substituting for $b_s = b_{s \setminus t} + b_t$ into the objective function reduces the problem to a m dimensional optimization problem in bids for the individual contracts b^i_s for $s \in S$. Assuming differentiability, the first order conditions are:

$$\sum_{t \text{ s.t. } s \subseteq t} G_t(b^i) + \sum_{w \subseteq S} \{ (b^i_w - c_w) \sum_{t \text{ s.t. } s \subseteq t} G^t_w(b^i) \} = 0 \quad \text{for all } s \in S$$

This is a system of m linear equations in $2^m - 1$ variables (the unobserved c_w). This system is under-identified. (iii) Follows from standard results for the single unit auction environments (see, e.g. Guerre et al., 2000). QED

Proof of proposition 4: The claim that any binding constraint introduces a single dimension of underidentification follows directly from (9) since the solution is unique up to the value of the Lagrangian multiplier involved in that constraint. The rest of the proof uses the following properties of determinants: (1) Determinants are invariant to linear transformations of rows or columns, (2) determinants are invariant to the permutation of rows and the corresponding column, (3)

$$\det \begin{bmatrix} a_{11} + b_{11} & .. & a_{1N} \\ .. & & .. \\ a_{N1} + b_{N1} & .. & a_{NN} \end{bmatrix} = \det \begin{bmatrix} a_{11} & .. & a_{1N} \\ .. & & .. \\ a_{N1} & .. & a_{NN} \end{bmatrix} + \det \begin{bmatrix} b_{11} & .. & a_{1N} \\ .. & & .. \\ b_{N1} & .. & a_{NN} \end{bmatrix}, \text{ and (3)}$$

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

Let G denote the $2^m - 1$ x 1 vector of G_s, define e_t as the $2^m - 1$ x 1 vector with entry 1 at the row corresponding to contract t and zero elsewhere and $I_{r = t \cup w}$ as the $2^m - 1$ x 1 vector with entry 1 in the row corresponding to contract r and -1 in the rows corresponding to contracts t and w. With these notations,

$$D(b^i; \mu, \lambda) = -G + \sum_{t \subseteq S} \mu_t e_t + \sum_r \sum_{\substack{t, w \subset r, \, t \cap w = \emptyset, \\ |t| \leq |w|}} \lambda_{r = t \cup w} I_{r = t \cup w}$$

Let $A_s B$ denote matrix A from which the column corresponding to contract s has been replaced by vector B. Cramer's rule together with properties (2) and (3) of determinants

38

imply that

$$
\begin{aligned}
b_s^{eff} - c_s &= \frac{1}{\det \nabla G} \det \nabla G_s D(b^i; \mu, \lambda) \\
&= \frac{1}{\det \nabla G}[-\det \nabla G_s G + \sum_{t \subseteq S} \mu_t \det \nabla G_s e_t \\
&\quad + \sum_{r \subseteq S} \sum_{\substack{t, w \subset r, \ t \cap w = \emptyset, \\ |t| \leq |w|}} \lambda_{r = t \cup w} \det \nabla G_s I_{r = t \cup w}]
\end{aligned}
$$
(15)

In words, the underlying cost parameter c_s depends linearly on the values of the Lagrangian multipliers.

Claim 1: $\det \nabla G_s e_t > 0$ if $t = s$, and zero otherwise.

Proof: If the reserve price on contract t is binding or if b_t^i is irrelevant (which is the only time we need to worry about the sign of $\det \nabla G_s e_t$), then $G_t^s(b_t^{eff}, b_{-s}^i) = 0$ for all $s \neq t$, and $G_t^t(b_t^{eff}, b_{-t}^i) < 0$, that is, the column in matrix ∇G that corresponds to contract t is all zeros but for the (row) entry corresponding to the tth contract. Now, by construction, the column in matrix $\nabla G_s e_t$ that corresponds to contract s is all zeros but for the (row) entry corresponding to the tth contract. If $s \neq t$, matrix $\nabla G_s e_t$ has two linearly dependent columns (corresponding to contracts s and t) so $\det \nabla G_s e_t = 0$. If $s = t$, $\det \nabla G_t e_t$ is equal to the determinant of matrix ∇G from which the row and column corresponding to contract t have been removed. By proposition 3, $\det \nabla G_t e_t > 0$ (since 2^m rows remain).

Claim 2: $\det \nabla G_s I_{r = t \cup w} > 0$ if $s = r$, it is < 0 if $s = t$ or w.

Proof: Suppose first that $r = s$. Define L_s as the operator that adds the values associated to row s to rows t and w so that $L_s \nabla G_s I_{s = t \cup w}$ becomes a matrix with a zero column at position s except for the "1" entry at row s. Define M as the 2^m x 2^m matrix made of matrix $L_s \nabla G_s I_{s = t \cup w}$ from which the row and the column corresponding to contract s have been removed, and M' as the 2^m x 2^m matrix made of matrix $L_s \nabla G$ from which the row and column corresponding to contract s have been removed. Then $M = M'$ and $\det M' = \det M > 0$ by lemma 1 and property (1). The claim follows from the fact that $\det \nabla G_s I_{s = t \cup w} = \det L_s \nabla G_s I_{s = t \cup w} = \det M$.

Now suppose that $s = t$ or w, and define L_s as the operator that adds row s to row r and substracts row s from row w so that $L_s(\nabla G_s I_{r = s \cup w})$ is now a matrix with a zero column except for an entry -1 at row s. Define M and M' as above. The claim follows from the fact that $\det \nabla G_s I_{r = s \cup w} = -\det M < 0$.

Proposition 4 then follows from (15), claims 1 and 2, and the fact that $\det \nabla G < 0$. QED

39

Proof of corollary 2: Suppose that $b_s^i = b_t^i + b_w^i$ for some t, w and s with $s = t \cup w$ and $t \cap w = \emptyset$. Define $\alpha_{t \cup w} = c_t + c_w - c_{t \cup w}$, the cost synergy between contracts t and w. With this notation, $b_s^i - c_s = b_t^i - c_t + b_w^i - c_w + \alpha_{t \cup w}$. Let C_s the operator on the columns of ∇G that add column s to columns t and w. We have:

$$\nabla G(b^i)[b^{eff} - c] = C_s \nabla G(b^i) \begin{bmatrix} \cdots \\ b_t - c_t \\ b_w - c_w \\ \alpha_{t \cup w} \\ \cdots \end{bmatrix} = D(b^i; \mu, \lambda)$$

Given that such transformation on the column of ∇G do not change the value of the determinant, we can use the same type of arguments as in the proof of proposition 4 (in particular, note that $\alpha_{t \cup w}$ is at the position corresponding to contract s) to prove that $\alpha_{t \cup w}$ depends negatively on the value of $\lambda_{\{t \cup w\} = t \cup w}$. Therefore, setting $\lambda_{\{t \cup w\} = t \cup w} = 0$ provides an upper bound to the synergy $\alpha_{t \cup w}$. QED

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, Susan and Phil Haile (2001), Identification of Standard Auction Models, forth. Econometrica

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

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

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

[8] Cohen A. C. and B. Jones Whitten (1980), Estimation in the Three-Parameter Log-normal Distribution, Journal of the American Statistical Association, 75, 399-404.

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

[10] Donald, S., H. Paarsch and J. Robert (2001), Identification and Estimation in Empirical Models of Sequential, Oral, Ascending Price Auctions, mimeo.

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

[12] Glaister, S. and M. Beesley (1991), Bidding for Tendered Bus Routes in London, Transportation Planning and Technology, 15, 349-366.

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

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

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

[16] Kennedy, D., S. Glaister and T. Travers (1995), London Bus Tendering, The Greater London Group at the London School of Economics and Political Science.

[17] Jofre-Bonet M. and M. Pesendorfer (2001), Estimation of a Dynamic Auction Game, mimeo, Yale University.

[18] London Transport Buses (1999), The Bus Tendering Process, available at http://www.londontransport.co.uk/buses/images/tend_rpt.pdf

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

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

[21] Li, Tong, Isabelle Perrigne and Quang Vuong (2000), Structural Estimation of the Affiliated Private Value Auction Model, University of Southern California, mimeo.

[22] 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.

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

[24] 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.

[25] Myerson, R. (1981), Optimal Auction Design, Mathematics of Operation Research, 6, 58-73.

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

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