Estimation of Multi-unit Double Auctions

Matheus Silva*

Advisor: Quang Vuong

Fall 2020

Abstract

I develop a resampling strategy to estimate private values in a multi-unit double auction, that is, auctions where many sellers and many buyers send schedules to a market organizer. This methodology allows the econometrician to estimate oligopolistic and oligopsonist market power jointly, and I apply it to Italian electricity market data. My results indicate that there is evidence both consumers and sellers exert market power to some extent.

1 Introduction

In this paper, I present a methodology to estimate two-sided auctions of multiple goods. In those auctions there is a finite number of sellers and buyers, each of whom submit schedules of quantities they are willing to negotiate depending on the price. A central market authority then sets a unique price that equilibrates supply and demand.

To do so, I extend Wilson's (1979) share auction model with multiple buyers to multiple sellers and a unique equilibrium price, as opposed to the pay-as-you-go pricing in his paper. The estimation procedure is based on Guerre, Perrigne, and Vuong (2000) and Hortaçsu (2002), which consists of nonparametrically estimating the first-order conditions of the agents' maximization problems. Using a simple solution to a particular case of this model, I assess the performance of the estimation algorithm using Monte Carlo simulations.

I also provide an application to electricity markets. One of the policy goals of those market makers is to minimize the room participants may have to manipulate the equilibrium price in their favor. The auction model in this paper provides a simple way to quantify market power: how far the true valuations are from the submitted price schedules. My estimation shows that the individuals I analyze exert market power by masking their true valuations, manipulating the equilibrium price in their favor.

The remainder of the paper is organized as follows: I describe the model in Section 2, then discuss identification and describe the estimation procedure in Section 3. Section 4 is devoted to Monte Carlo simulations to assess the efficacy of the procedure. Finally, in Section 5, I focus on the empirical application, describing the structure of the day-ahead electricity market in Italy and presenting the results of my estimation.

*I would like to thank all participants of the third-year paper seminars for their questions comments.

1.1 Literature review

Optimal nonparametric estimation of auctions date back to Guerre, Perrigne and Vuong (2000). Their approach, widely used in the applied literature, consists of estimating nonparametrically the first-order conditions of the agent's expected utility maximization problem.

Since then, the literature was developed by adding other features of the data into an auction model and, subsequently, its estimation. Despite its empirical relevance, the literature on the estimation of multi-unit auctions still has many gaps. On one hand, one-sided multi-unit auctions have been studied deeply (particularly in the context of treasury auctions); see Hortaçsu (2001, 2002), Hortaçsu and McAdams (2010), Kastl (2011), and McAdams (2008). On the other, two-sided multi-unit auctions have not.

While there is theoretical work on two-sided multi-unit auctions (for example, McAdams (2006) studies the characteristics of their equilibria), their estimation remains somewhat unexplored as Hortaçsu and McAdams (2018) show. My paper contributes by proposing a resampling methodology to estimate private marginal valuations in this framework.

When it comes to my empirical application, I am focused on the estimation of market power in electricity markets. Decentralized and self-regulated electricity markets typically have day-ahead operations that can be modelled as a two-sided multi-unit auction. From my research, much of the work in this area closely follows Wolak (2002, 2003), who proposes estimating market power with suppliers' bid data by imposing a functional form on the cost structure of the generators and estimating that model by GMM. Subsequent empirical work consists of applications of this methodology to the specific markets the researchers are interest in. In my interpretation, this approach faces two severe shortcomings: the first is that it requires imposing a specific functional form on the cost function of the generators since it relies on GMM estimation; the second is that they take into consideration only the supplier side, which I interpret as taking the demand as fixed.

Alternative methodologies tackle in particular the first problem. Reguant (2014) and Mar and Reguant (2013) use of nonparametric auction estimation procedures to study the Spanish electricity market; Rossetto, Grossi, and Pollitt (2019) apply a synthetic supply approach to study the Italian market as well. All of these papers cited so far are particularly concerned about the market power exerted by suppliers, but it is also a concern that buyers might have the ability to manipulate prices as well, especially given their size. Bigerna and Bollino (2016) explore this dimension by measuring unilateral market power in the demand side of the Italian electricity market.

My contribution here is to tackle both problems (imposing of a functional form and fixing one side of the market) simultaneously. The reasons for this is that I estimate the first-order conditions of all agents nonparametrically and, later in Section 3, my proposed bootstrapping procedure depends on variation on both sides.

2 The model

The model is based on Wilson (1979), Hortaçsu (2002), and Hortaçsu and McAdams (2011). The market is comprised of two types of agents, M buyers and N sellers, who simultaneously send bid and ask schedules to a market maker. The good in question is perfectly divisible. These functions are denoted y and x, respectively. The market authority, in turn, aggregates the individual demand and supply schedules and calculates the equilibrium price, which all individuals pay[1]. Market power manifests itself here as the difference between the schedules sent to the auctioneer and the agents' private marginal valuations.

Nature assigns a private signal for each agent. For ease of notation, let $s_i, i = 1, ..., M$ be the private signal of the i^{th} buyer and $\psi_j, j = 1, ..., N$ be the private value of the j^{th} seller. I work under the assumption that all s_i and ψ_j are mutually independent, but they can follow different absolutely continuous distributions, $s_i \sim F(\cdot)$ and $\psi_j \sim \tilde{F}(\cdot)$. The marginal valuations are given by the functions ν_i (for a buyer) and ξ_j (for a seller.) In the most general setup, ν_i and ξ_j can depend not only on the quantity of the good and the individuals' private signals, but also on everyone else's signals. The IPV setup restricts ν_i and ξ_j to be the same across individuals and not on others' valuations[2]. In other words, $\nu_i(q, s_i, s_{-i}) = \nu(q, s_i)$, $\forall i$ and $\xi_j(q, \psi_j, \psi_{-j}) = \xi(q, \psi_j)$, $\forall j$.

I will now define the distribution of equilibrium prices that is the central object of this paper. The equilibrium price is determined by the intersection of supply and demand,

$$\sum_{i=1}^{M} y_i(p^c) = \sum_{j=1}^{N} x_j(p^c)$$

And the probability distribution of interest is:

$$G(p) = \mathbb{P}(p^c \leq p)$$

This is key to this model given that all individuals will maximize their expected utility according to this probability measure. However, given that each player knows their private valuation, buyers will work with the following distribution:

$$H(p, y_i(p)) = \mathbb{P}(p^c \leq p | y_i(p))$$

Which can be rewritten in terms of the residual supply,

$$H(p, y_i(p)) = \mathbb{P}(y_i(p) \leq RS_{-i}(p) | y_i(p)), \text{where } RS_{-i}(p) \equiv \sum_{j=1}^{N} x_j(p) - \sum_{k \neq i}^{M} y_k(p)$$

Similarly, a seller will work with

$$\tilde{H}(p, x_j(p)) = \mathbb{P}(x_j(p) \leq RD_{-j}(p) | x_j(p)), \text{where } RD_{-j}(p) \equiv \sum_{i=1}^{M} y_i(p) - \sum_{k \neq j}^{N} x_k(p)$$

Given this, the problem a buyer solves is:

$$\max_{\{y_i(\cdot)\}} \mathbb{E}U_i(s_i) = \int_0^{\infty} \left[\int_0^{y_i(p)} \nu_i(u) du - p y_i(p) \right] dH(p, y_i(p))$$

The first-order condition for this problem is:

$$\nu(y(p, s_i), s_i) = p - y(p, s_i) \frac{\frac{\partial H}{\partial y}(p, y(p, s_i))}{\frac{\partial H}{\partial p}(p, y(p, s_i))}$$

The problem a seller solves is:

$$\max_{\{x_j(\cdot)\}} \mathbb{E}\Pi_j(\psi_j) = \int_0^{\infty} \left[p x_j(p) - \int_0^{x_j(p)} \xi_j(u) du \right] d\tilde{H}(p, x_j(p))$$

[1] This contrasts with Hortaçsu's (2002) discriminatory pricing setting.

[2] Wilson (1979) is a model where agents have affiliated values. A two-sided version of that model has still to be developed.

In this case, the first-order condition is:

$$\xi(x(p,\psi_j),\psi_j) = p + x(p,\psi_j)\frac{\frac{\partial \tilde{H}}{\partial x}(p, x(p,\psi_j))}{\frac{\partial \tilde{H}}{\partial p}(p, x(p,\psi_j))}$$

See the Appendix for the complete derivation of these equations.

It is also important to state that in multi-unit auctions like this one, the equilibrium strategies are monotonic, which is proven in McAdams (2006). In the next section, I discuss how important this is to the identification of the model.

Another fact worth mentioning is that, even though the schedules in the model I solved are continuous functions of the price, this is not what one sees in the data. Typically, agents submit step functions: a set of K pairs $\{(y_k, p_k)\}_{k=1}^{K}$ that indicate their willingness to buy or sell the auctioned good. I will carry this notation throughout the remainder of this paper, but I will not take this into consideration in my estimations. For a more comprehensive view on this problem, see Kastl (2011).

3 Identification

To achieve identification of the ν_i and ξ_j functions, the econometrician must be able to estimate the distribution of equilibrium prices from the data. In simple cases, where the equilibrium is given by the intersection between the supply and demand curves, this problem boils down to nonparametrically estimating the $\eta_i(p,q)$ probability distribution function,

$$\eta_i(p,q) = \mathbb{P}\left[q \leq RS_{-i}(p)\right]$$

Which is analogous to the H function presented before[3], i.e., the probability that the quantity q is less than the random residual supply faced by consumer i.

Following Elyakime et al (1994), Guerre, Perrigne, and Vuong (2000), and Hortaçsu (2002), the econometrician can estimate $\eta_i \forall (p,q)$ if they can estimate the joint distribution of $\{y_k(p,s_k), k \neq i; x_j(p,\psi_j)\}$ from the data.

In more complex cases where the welfare maximization problem faces restrictions, for example, identification will rely on being able to estimate the distribution of equilibrium prices and quantities along the joint distribution of $\{y_k(p,s_k), k \neq i; x_j(p,\psi_j)\}$. Meaning the econometrician not only needs the data on the schedules themselves, but must also be able to replicate the full process of market clearing.

On top of these, identification of the model requires all bids to be the unique solution to the optimization problem described before and monotonic functions of prices. The end of Section 2 presents a brief discussion of these issues.

Finally, if one can estimate η_i from the data, then evaluating it at the individual's demand function backs out an estimate for $H(p, y_i(p))$. The partial derivatives needed for estimation are also identified without any extra costs, since probability laws require these to integrate to 1.

$$\begin{aligned} H(p, y_i(p)) = \eta_i(p,q)|_{q=y_i(p)} &\rightarrow \frac{\partial H(p, y_i(p))}{\partial p} = \frac{\partial \eta_i(p,q)|_{q=y_i(p)}}{\partial p} \\ &\rightarrow \frac{\partial H(p, y_i(p))}{\partial y} = \frac{\partial \eta_i(p,q)|_{q=y_i(p)}}{\partial y} \end{aligned}$$

[3] In the case of the supplier, it is enough to redefine the variables.

From the perspective of an applied econometrician, identification will "come for the data" in the sense they need to check whether the bids and asks are monotonic in prices. This does not pose a problem in many applications since the auctioneer may require the submitted schedules to be monotonic, which is the case in the application provided in this paper. A more sensitive point is that the model will only be identified if the econometrician can perfectly replicate the market clearing procedure, that is, with the same information as the auctioneer, the econometrician must be able to replicate the market equilibrium. This is more complicated in practice, since many auctioneers may not disclose the entirety of their market clearing algorithm[4].

It is also important to notice that this argument implies that $\nu_i(p)$ and $\xi_j(p)$ are identified. In other words, the econometrician can only identify the functions ν and ξ jointly with the private signals, and not separately. As Perrigne and Vuong (2020) argue, identifying ν or ξ separately from the individual signals requires two extra restrictions:

(i) The signals are uniformly distributed; and,

(ii) The schedules submitted by the agents are increasing functions of the private values.

Under these conditions, the econometrician can apply the procedures described in Matzkin (2003). The underlying reason for this is that now the functions ν and ξ are held constant and restrictions (i) and (ii) tell the econometrician how they behave in the private signal. Therefore, separate identification is now possible. While this implies some loss in flexibility, it opens an avenue for other comparative exercises. For example, one could address questions like "What would happen if a smaller electricity company was replaced by a bigger one, with more generators?" or "What happens if one client leaves the market?" This is a venue I will explore in further work, and I will reiterate this point later in the conclusion of this paper.

4 Estimation

One can estimate the model presented in Section 3 in three different ways. In this version of the paper, I explore two avenues. First I estimate it using a kernel estimator à la Guerre et al (2000) and then a bootstrap-type estimator as in Hortacsu (2001,2002) and Hortaçsu and McAdams (????? 2011). The estimator I leave out for now uses Matzkin's (2003) estimation strategy, I discuss its advantages later in this Section.

I present each of these estimation strategies within the framework provided in Section 3, which means we don't require any sort of conditioning and adjustments due to the simple, known rules of the auction. However, the practitioner should be aware of these simplifications and incorporate the relevant variables for market clearing in their estimation. For example, Perrigne and Vuong (2020) argue that Hortaçsu's (2002) estimation of his share auction model should take into consideration the total quantity of the auctioned good available to the public.

4.1 Kernel estimator

The first approach is to estimate the partial derivatives in the first-order condition directly with a kernel estimator.

In the case of a buyer [the probability distribution they face is:]

[4] Reguant (2014) relies on an approximation of the market clearing procedure in the Spanish electricity market, for example.

$$\hat{H}(p,y) = \frac{1}{(M+N)L}\sum_{\ell=1}^{L}\sum_{i=1}^{I}\tilde{K}\left(\frac{\sum_j x_{j\ell}(p) - y - \sum_{k\neq i}^{I} y_{k\ell}(p)}{h}\right)$$

Where $\tilde{K}(u) = \int_{-\infty}^{u} K(\nu)d\nu$ and h is the kernel bandwidth.

Taking the partial derivative with respect to y,

$$\hat{H}_y(p,y) = -\frac{1}{(M+N)Lh}\sum_{\ell=1}^{L}\sum_{i=1}^{I}K\left(\frac{\sum_j x_{j\ell}(p) - y - \sum_{k\neq i}^{I} y_{k\ell}(p)}{h}\right)$$

And with respect to p,

$$\hat{H}_p(p,y) = -\frac{1}{(M+N)Lh}\sum_{\ell=1}^{L}\sum_{i=1}^{I}\left(\sum_j x'_{j\ell}(p) - \sum_{k\neq i}^{I} y'_{k\ell}(p)\right)K\left(\frac{\sum_j x_{j\ell}(p) - y - \sum_{k\neq i}^{I} y_{k\ell}(p)}{h}\right)$$

Plugging the partial derivatives into the first-order condition, I obtain the estimator of the marginal value of a buyer.

$$\hat{\nu}(p_k, y_i(p_k)) = p_k - y_i(p_k)\frac{\hat{H}_y}{\hat{H}_p}$$

$$\hat{\nu}(p_k, y_i(p_k)) = p_k - y_i(p_k)\left[\frac{\sum_{\ell=1}^{L}\sum_{i=1}^{I}K\left(\frac{\sum_j x_{j\ell}(p) - y - \sum_{k\neq i}^{I} y_{k\ell}(p)}{h}\right)}{\sum_{\ell=1}^{L}\sum_{i=1}^{I}\left(\sum_j x'_{j\ell}(p) - \sum_{k\neq i}^{I} y'_{k\ell}(p)\right)K\left(\frac{\sum_j x_{j\ell}(p) - y - \sum_{k\neq i}^{I} y_{k\ell}(p)}{h}\right)}\right]$$

The estimator for the seller is symmetric, a direct consequence of the nature of the model. Explicitly,

$$\hat{\hat{H}}(p,x) = \frac{1}{(M+N)L}\sum_{\ell=1}^{L}\sum_{j=1}^{N}\tilde{K}\left(\frac{\sum_{i=1}^{I} y_{i\ell}(p) - \sum_{k\neq j} x_{k\ell}(p) - x}{h}\right)$$

Taking the partial derivative with respect to x and p and forming the estimate, I obtain:

$$\hat{\hat{H}}_x(p,x) = -\frac{1}{(M+N)Lh}\sum_{\ell=1}^{L}\sum_{i=1}^{I}K\left(\frac{\sum_{i=1}^{I} y_{i\ell}(p) - \sum_{k\neq j} x_{k\ell}(p) - x}{h}\right)$$

And with respect to p,

$$\hat{\hat{H}}_p(p,x) = -\frac{1}{(M+N)Lh}\sum_{\ell=1}^{L}\sum_{i=1}^{I}\left(\sum_{i=1}^{I} y'_{i\ell}(p) - \sum_{k\neq j} x'_{k\ell}(p)\right)K\left(\frac{\sum_{i=1}^{I} y_{i\ell}(p) - \sum_{k\neq j} x_{k\ell}(p) - x}{h}\right)$$

With which we can form the estimate of the marginal value of each supplier,

$$\hat{\xi}(p_k, x_j(p_k)) = p_k + x_j(p_k)\left[\frac{\sum_{\ell=1}^{L}\sum_{i=1}^{I}K\left(\frac{\sum_{i=1}^{I} y_{i\ell}(p) - \sum_{k\neq j} x_{k\ell}(p) - x}{h}\right)}{\sum_{\ell=1}^{L}\sum_{i=1}^{I}\left(\sum_{i=1}^{I} y'_{i\ell}(p) - \sum_{k\neq j} x'_{k\ell}(p)\right)K\left(\frac{\sum_{i=1}^{I} y_{i\ell}(p) - \sum_{k\neq j} x_{k\ell}(p) - x}{h}\right)}\right]$$

4.2 A bootstrap estimator

I will estimate this model using an extended version of the bootstrap strategy proposed in Hortaçsu (2001) and Hortaçsu and McAdams (2011). This boils down to resampling the residual supplies or demands, fixing the individual of interest, and then using the bootstrapped H or $\tilde{H}$ distributions to compute the first-order conditions. The econometrician observes all bid and ask schedules for a total of L auctions. Algorithms 1 and 2 show how to estimate the private marginal values.

Algorithm 1. *(Bootstrap algorithm for the buyers)*

1. *Fix buyer i.*
2. *Repeat the following steps B times:*
   - *Resample N suppliers' and $M-1$ clients' schedules, excluding i.*
   - *Compute and store the residual supply, RS^b_{-i}.*
3. *Compute the partial derivatives H_p and H_y numerically, obtaining $\hat{H}_p$ and $\hat{H}_y$. (I show how to compute them numerically in the end of this section.)*
4. *Compute the estimated marginal value at each price p_k:*

$$\hat{\nu}(p_k, y_i(p_k)) = p_k - y_i(p_k)\frac{\hat{H}_y}{\hat{H}_p}$$

Algorithm 2. *(Bootstrap algorithm for the sellers)*

1. *Fix seller j.*
2. *Repeat the following steps B times:*
   - *Resample M clients' $N-1$ suppliers' schedules, excluding j.*
   - *Compute and store the residual demand, RD^b_{-j}.*
3. *Compute the partial derivatives $\tilde{H}_p$ and $\tilde{H}_y$ numerically, obtaining $\hat{\tilde{H}}_p$ and $\hat{\tilde{H}}_y$. (I show how to compute them numerically in the end of this section.)*
4. *Compute the estimated marginal value at each price p_k:*

$$\hat{\xi}(p_k, x_j(p_k)) = p_k + x_j(p_k)\frac{\hat{\tilde{H}}_x}{\hat{\tilde{H}}_p}$$

The following figure illustrates the bootstrap procedure for a given individual on the demand side. Once the residual supplies are resampled, I compute the intersection point between the curves. Multiple iterations of these steps will back out the H distribution, which is the main object of interest.

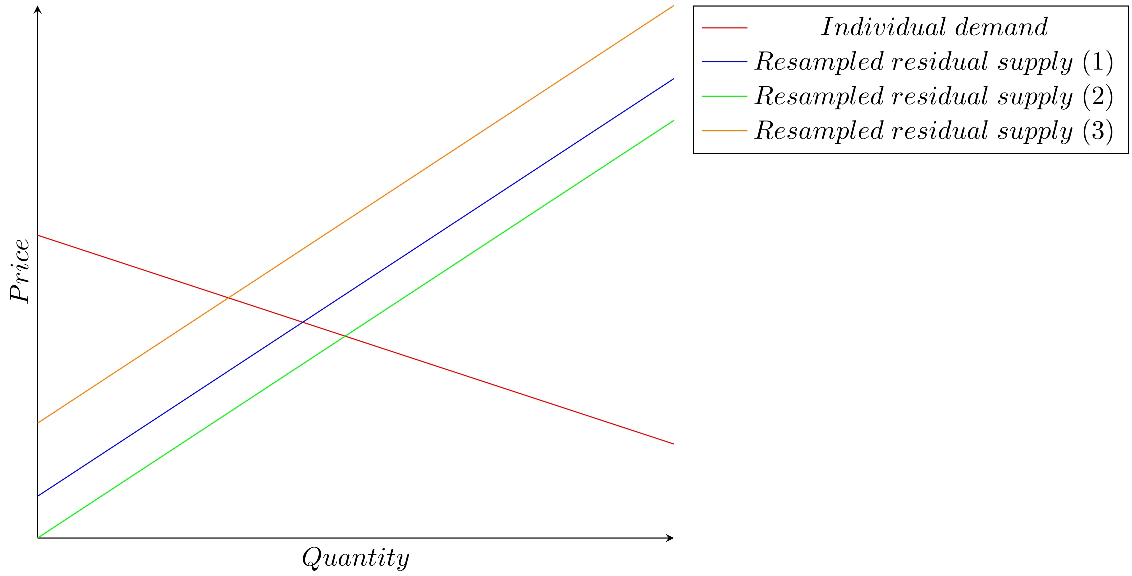


Figure 1: Illustration of the bootstrap procedure to estimate the marginal values.

Naturally, this estimation procedure depends on the consistency of this bootstrap procedure.

Proposition 3. *The bootstrap algorithm backs up the marginal valuations consistently.*

Proof. Given the discussion on identification, as the number of auctions L increases

$$\lim_{B\to\infty} \frac{1}{B}\sum_{i=1}^{B} \mathbb{I}[y_i(p_k) \leq RS_i^b(p_k)] \to H(p_k, y_i(p_k))$$

and

$$\lim_{B\to\infty} \frac{1}{B}\sum_{i=1}^{B} \mathbb{I}[x_j(p_k) \leq RD_j^b(p_k)] \to \tilde{H}(p_k, y_i(p_k))$$

A direct implication of the Glivencko-Cantelli theorem. This means that the estimates

$$\hat{\nu}(p_k, y_i(p_k)) = p_k - y_i(p_k)\frac{\hat{H}_y}{\hat{H}_p}$$

and

$$\hat{\xi}(p_k, x_j(p_k)) = p_k + x_j(p_k)\frac{\hat{\tilde{H}}_x}{\hat{\tilde{H}}_p}$$

Also converge to their populational values $\nu(p_k, y_i(p_k))$ and $\xi(p_k, x_j(p_k))$, respectively. □

To conclude this computation section, it remains to describe how to calculate the numerical partial derivatives. There are many ways to compute them and picking one will ultimately depend on the richness of the data available, since obtaining more precise numerical estimates comes at the expense of more steps. Here are three, popular alternatives:

1. *(Hortaçsu (2002), Hortaçsu and McAdams (2011))*[5]

   Algorithm:

[5] In essence, this will not be an actual partial derivative since when the intersection changes, both price and quantities change. If the dataset were comprised by schedules with steps really close to one another, then the numerical results would not be greatly affected by calculating the derivative in this way. However, the steps in the schedules (on both sides of the market) are far enough from each other to provide unreliable estimates when this procedure is applied.

- Save bootstrap equilibrium prices and quantities, then estimate their joint cdf using Silverman's bandwidth method.
- Obtain the numerical partial derivatives and plug them into the first-order condition.

From his notation, these derivatives boil down to:

$$\hat{H}_y = \frac{\hat{H}(p_{t+1}, y_i(p_{k+1})) - \hat{H}(p_k, y_i(p_k))}{y_i(p_{k+1}) - y_i(p_k)}$$
$$\hat{H}_p = \frac{\hat{H}(p_{k+1}, y_i(p_{k+1})) - \hat{H}(p_k, y_i(p_k))}{p_{k+1} - p_k}$$

Therefore,

$$\frac{\hat{H}_y}{\hat{H}_p} = \frac{p_k - p_{k-1}}{y_i(p_k) - y_i(p_{k-1})}$$

2. *(Backwards derivative)*

$$\hat{H}_y = \frac{\hat{H}(p_k, y_i(p_k)) - \hat{H}(p_k, y_i(p_{k-1}))}{y_i(p_k) - y_i(p_{k-1})}$$
$$\hat{H}_p = \frac{\hat{H}(p_k, y_i(p_k)) - \hat{H}(p_{k-1}, y_i(p_k))}{p_k - p_{k-1}}$$

3. *(Two-sided numerical derivative)*

$$\hat{H}_y = \frac{\hat{H}(p_k, y_i(p_{k+1})) - \hat{H}(p_k, y_i(p_{k-1}))}{y_i(p_{k+1}) - y_i(p_{k-1})}$$
$$\hat{H}_p = \frac{\hat{H}(p_{k+1}, y_i(p_k)) - \hat{H}(p_{k-1}, y_i(p_k))}{p_{k+1} - p_{k-1}}$$

4.3 Matzkin

TBD

4.4 Discussion

Each of these methods have their advantages.

5 Monte Carlo simulations

To provide proper Monte Carlo simulations, I must first obtain a solution to the model. This involves working with the probability distributions H and $\tilde{H}$, which are complicated objects to deal with analytically. A simple case involves guessing that linear true marginal valuations on both the demand and the supply sides generate linear bid and ask functions, and verifying it is true under some conditions[6].

Assume that the true demand function of a buyer is:

$$D(p, s_i) = \alpha + \beta s_i + \gamma p, \ \alpha > 0, \gamma < 0$$

The inverse demand gives us the true valuation of the individual:

$$\nu(q, s_i) = \frac{1}{\gamma}(q - \alpha - \beta s_i)$$

[6] I follow Hortaçsu's (2001) approach to obtain this solution.

On the supply side, things work out similarly. The true supply function is:

$$S(p, \psi_j) = \tilde{\alpha} + \tilde{\beta}\psi_j + \tilde{\gamma}p,\ \tilde{\alpha} > 0, \tilde{\gamma} > 0$$

The inverse supply gives the true valuation of the supplier:

$$\xi(q, \psi_j) = \frac{1}{\tilde{\gamma}}(q - \tilde{\alpha} - \tilde{\beta}\psi_j)$$

The guesses are that the bid and ask functions, determined by the true parameters in the model, are linear in the private signals and prices. That is,

$$y(p, s_i) = a + bs_i + cp$$
$$x(p, \psi_j) = \tilde{a} + \tilde{b}\psi_j + \tilde{c}p$$

Proposition 4. *If*

(i) c_i, ψ_j *are mutually independent* $\forall i, j$,

(ii) $c_i, \psi_j \sim Exponential(\lambda), \forall i, j$,

then the model admits a linear solution as described above with $b = -\tilde{b}$.

Proof. See appendix for a complete proof. ☐

I use the following parametrization of the model for the simulations:

Parameter	Interpretation	Value
α	Parameter of the true demand marginal value function	0.9643
β	Parameter of the true demand marginal value function	-1.0714
γ	Parameter of the true demand marginal value function	-0.9286
$\tilde{\alpha}$	Parameter of the true supply marginal value function	0.0489
$\tilde{\beta}$	Parameter of the true supply marginal value function	0.9781
$\tilde{\gamma}$	Parameter of the true supply marginal value function	0.2781
N	Number of suppliers	5
M	Number of clients	30

Table 1: Set of structural parameters used in the Monte Carlo simulations.

And this set of parameters yields the following coefficients:

Parameter	Interpretation	Value
a	Solution of the intercept of the bid function	0.9
b	Solution of the coefficient associated with s_i of the bid function	-1
c	Solution of the coefficient associated with p of the bid function	-1
$\tilde{a}$	Solution of the intercept of the ask function	0.05
$\tilde{b}$	Solution of the coefficient associated with ψ_j of the ask function	1
$\tilde{c}$	Solution of the coefficient associated with p of the ask function	0.3

Table 2: Solution of the model implied by the set of parameters presented earlier.

Figure 2 presents the results of the estimation of the marginal values of one buyer using the kernel-based estimator presented in Section 4.1. It is noticeable that the estimator performs well in this experiment, mostly due to the simplicity of the setup and the sample size.

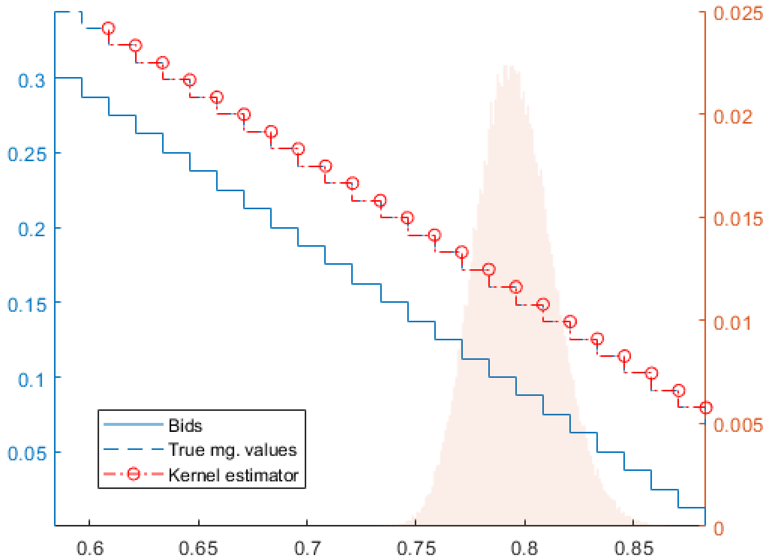


Figure 2: Results of the estimation for a single buyer, using the kernel-based estimator.

Figures 3 and 4 present the results of the Monte Carlo simulations using the resampling estimator for a single buyer. Figure 3 includes the estimates obtained with the one-sided backwards and the two-sided numerical derivatives presented previously. It's worth noticing that the estimated curves come to close to the actual marginal valuation, with the two-sided derivative performing better at the expense of an extra step of the submitted demand curve[7]. Another important thing to keep track of is how the quality of the estimates depends on the distribution of equilibrium prices and quantities (the histogram of the latter is graphed in red[8]). This is particularly noticeable in the region of the histogram closer to its center of mass.

[7]In fact, one could refine further the estimate of the numerical derivative if they're willing to pay the price in terms of observed steps.

[8]This suffices since the model is entirely linear.

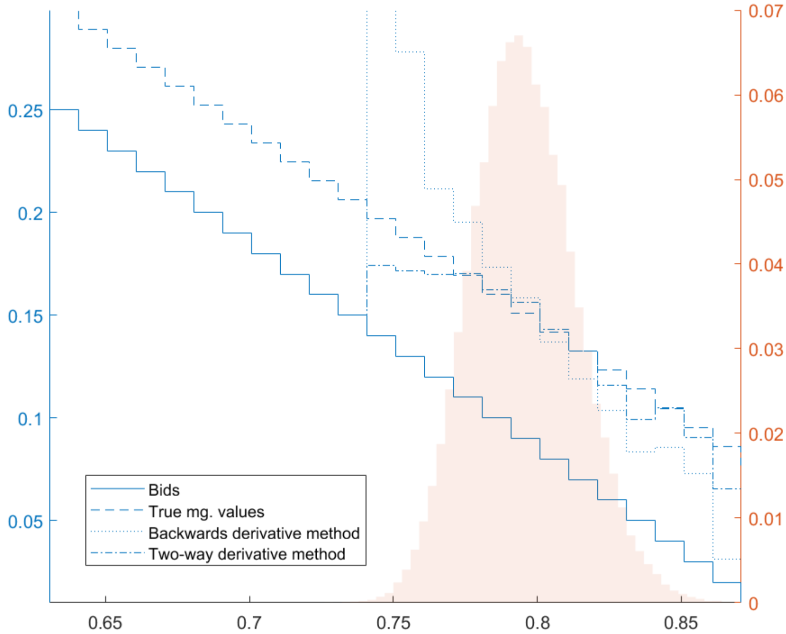


Figure 3: Results of the estimation for a single buyer, excluding Hortaçsu's (2002) method.

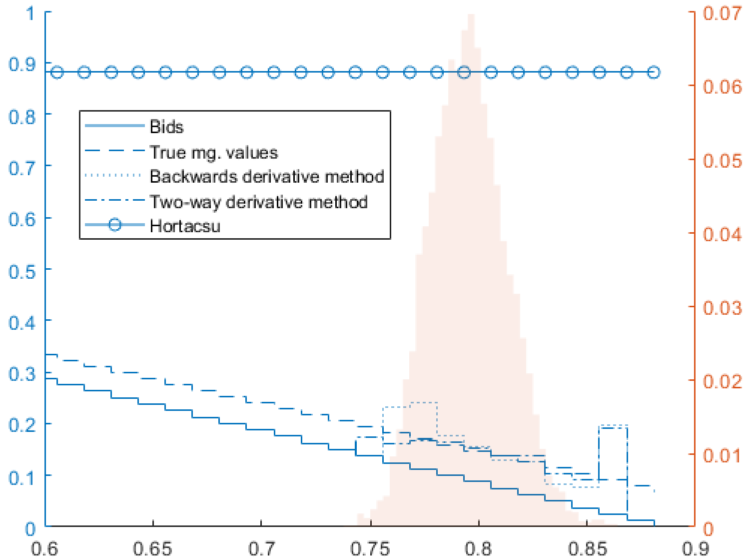


Figure 4: Results of the estimation for a single buyer, including Hortaçsu's (2002) method.

The numerical derivative described in Hortaçsu (2002) and Hortaçsu and McAdams (2011) yields a constant marginal valuation across quantities. In this example, this stems from the fact that the bid functions are linear. I present the results of the same Monte Carlo exercise in Hortaçsu's (2001, 2002) environment in the Appendix.

6 An empirical application: Market power in the electricity market

Parts of decentralized wholesale electricity markets operate in a double auction format. In the day-ahead market, agents sell and buy electricity for the next day by sending supply and demand schedules to a centralized market operator, who aggregates all this information and sets an equilibrium price.

The goal here is to measure market power in the sense provided by the model developed in Section 2: the ability agents have to mask their true valuations to increase their surpluses. There are different ways to approach this issue. One of the firsts is Wolfram (1999), who runs a regression on British data. A more influential one is Wolak (2002), who estimates a version of the first order condition presented in Section 2. Other papers include Fabra and Reguant (2013), Reguant (2014), and for the Italian market in particular, Bigerna and Bollino (2016). The analysis provided by my model differs from those other papers in this literature in two ways: first, a common approach is to impose a functional form on the cost functions of the firms; second, many analyses study only one side of the market, disregarding potential interactions between oligopolists and oligopsonists. The approach presented in this paper avoids both pitfalls by not imposing a functional form and, because of the nonparametric nature of the bootstrap procedure, working with both sides of the market simultaneously.

6.1 Overview of the day-ahead electricity market in Italy

The Italian electricity market is one of the largest in Europe. Not only it is a platform to negotiate the commodity within the country, but it is also connected with other countries such as France, Austria, Italy, Greece, Slovenia, Malta, and Switzerland. Due to demand expanding, it has seen an increase in negotiated quantities over the previous years. In 2018, the total amount sold was 567 TWh.

As in many other decentralized markets, the GME (*Gestore dei Mercati Energetici*, the market maker) organizes the markets in two ways: the first is the day-ahead market (DAM for short) and the real-time market. The former is the focus of this paper. In the day-ahead market, agents submit their schedules and the resulting equilibrium is taken as a base for further negotiations in the real-time market[9].

In contrast to other electricity market operators, the GME institutes a unique national price, the *Prezzo Unico Nazionale* (henceforth PUN). Which maps this problem directly into the framework developed in Section 2.

A comprehensive review on many electricity markets around the world and how they compare in size and operation can be found in Shah and Chatterjee (2020).

The day-ahead market is composed by buyers and sellers that agree on purchasing and delivering energy for the next day (relative to the day the market closes) according to schedules they submit, independently and anonymously, to the GME. The market operator then clears the market by finding a unique price that maximizes aggregate welfare. Timing-wise, the GME starts taking offers for a specific day nine days in advance, starting at 8 a.m. The markets close at 12 p.m. the day prior to the delivery.

In principle, the procedure to find the equilibrium point would simply consist of crossing the supply and demand curves. However, the electricity market faces physical constraints since the system has a certain transmission capacity. The Italian market also categorizes individuals according to some priority rating (merit-order restriction). The wealth of the data provided by the GME allows the econometrician to replicate the

[9]A parallel to this is the stock market. The opening price of a stock in a given day is a result of a two-sided multi-unit auction conducted in the after-hour market the night preceding the beginning of that day.

equilibrium price. In subsection 6.3, I describe the procedure in more depth.

6.2 Description of the dataset

The GME makes a rich dataset on the day-ahead market publicly available and fully downloadable from their website[10]. For every hour of each day, the econometrician can observe all bids and ask schedules, each merit-order submission, and each order involving other countries. Moreover, the location of each individual generator and each potential client is observed. As usual in electricity markets, these are important to clear the zonal markets. Finally, the last set of relevant information comprehends the physical capacity of the transmission lines across regions. This is relevant since there might exist congestion in the systems, which affects the equilibrium price by creating scarcity.

I also define a market as a one-hour window since individuals cannot submit block orders, thus making each hour window separate from each other. The sample used for the estimation goes from July 1^{st} to July 31^{st}, 2019. I will later focus the discussion on the 12-1 a.m. window market. The histograms in Figure 5 and Table 5 describe the distribution of market clearing prices in each one-hour market segment. Moreover, tables 3 and 4 on the following pages summarize both the demand side[11] and the supply curves. Since the GME collects step functions, I highlight the statistics about the number of steps as well. Altogether, these show significant variation in prices within each hour of the day and that the amount of agents is small, justifying the concern about market power on both sides of the market.

[10] For example, one can access data from January 27^{th}, 2018 through these links:

https://www.mercatoelettrico.org/en/Download/DownloadDati.aspx?val=OfferteFree_Pubbliche

https://www.mercatoelettrico.org/It/WebServerDataStore/MGP_LimitiTransito/20180127MGPLimitiTransito.xml

https://www.mercatoelettrico.org/It/WebServerDataStore/MGP_Prezzi/20180127MGPPrezzi.xml

https://www.mercatoelettrico.org/It/WebServerDataStore/MGP_Quantita/20180127MGPQuantita.xml

[11] For the clients that submit bid functions, as opposed to those that submit must-take orders.

Demand side				
Hour	Average number of clients	Std. Dev	Average number of steps	Std. Dev
01	8.29	0.46	3.90	0.60
02	8.29	0.46	3.65	0.58
03	8.29	0.46	3.55	0.50
04	8.26	0.44	3.53	0.46
05	8.26	0.44	3.57	0.44
06	8.26	0.44	3.65	0.54
07	8.23	0.43	3.89	0.55
08	8.23	0.43	3.91	0.59
09	8.23	0.43	3.71	0.67
10	8.23	0.43	3.73	0.60
11	8.23	0.43	3.86	0.55
12	8.16	0.45	3.82	0.46
13	8.23	0.50	3.85	0.52
14	8.26	0.51	3.80	0.56
15	8.16	0.45	3.73	0.47
16	8.16	0.45	3.69	0.51
17	8.16	0.45	3.72	0.45
18	8.19	0.40	3.74	0.46
19	8.16	0.37	3.65	0.54
20	8.16	0.37	3.55	0.59
21	8.16	0.37	3.50	0.61
22	8.16	0.37	3.57	0.55
23	8.16	0.37	3.95	0.55
24	8.16	0.37	3.84	0.51

Table 3: Descriptive statistics of the buyers' side, per hour. **Source:** GME.

Supply Side				
Hour	Average number of participants	Std. Dev.	Average number of steps	Std. Dev.
01	13.06	0.25	115.71	4.70
02	12.97	0.48	114.91	5.78
03	13.03	0.55	114.41	6.15
04	13.06	0.51	114.58	6.54
05	13.06	0.51	114.88	6.36
06	13.06	0.51	119.55	6.31
07	12.94	0.36	130.07	5.83
08	12.81	0.40	135.15	7.15
09	12.68	0.60	136.45	9.42
10	12.58	0.62	138.17	10.33
11	12.52	0.63	138.73	10.18
12	12.45	0.62	139.20	9.51
13	12.39	0.62	140.53	9.83
14	12.39	0.62	141.39	10.03
15	12.42	0.62	140.69	9.80
16	12.52	0.63	140.36	10.16
17	12.48	0.63	141.72	10.20
18	12.55	0.62	142.06	10.17
19	12.39	0.72	145.35	10.87
20	12.45	0.62	142.45	8.92
21	12.26	0.68	138.71	9.33
22	12.26	0.68	127.62	8.09
23	12.19	0.70	128.28	7.86
24	12.13	0.67	129.15	8.01

Table 4: Descriptive statistics of the suppliers' side, per hour. **Source:** GME.

Prices		
Hour	Average price	Std. Dev.
01	51.03	6.08
02	48.59	5.77
03	45.89	4.84
04	44.47	4.49
05	43.71	4.31
06	43.85	4.44
07	46.20	6.62
08	50.26	8.13
09	54.24	9.22
10	54.98	9.40
11	53.44	8.75
12	52.38	9.24
13	48.57	7.21
14	48.16	7.72
15	51.12	9.86
16	53.47	11.28
17	55.55	11.80
18	57.21	10.78
19	58.64	9.68
20	61.74	9.25
21	62.10	7.94
22	62.44	7.60
23	56.99	6.41
24	50.48	5.62

Table 5: Descriptive statistics of the market clearing prices, per hour. **Source:** GME.

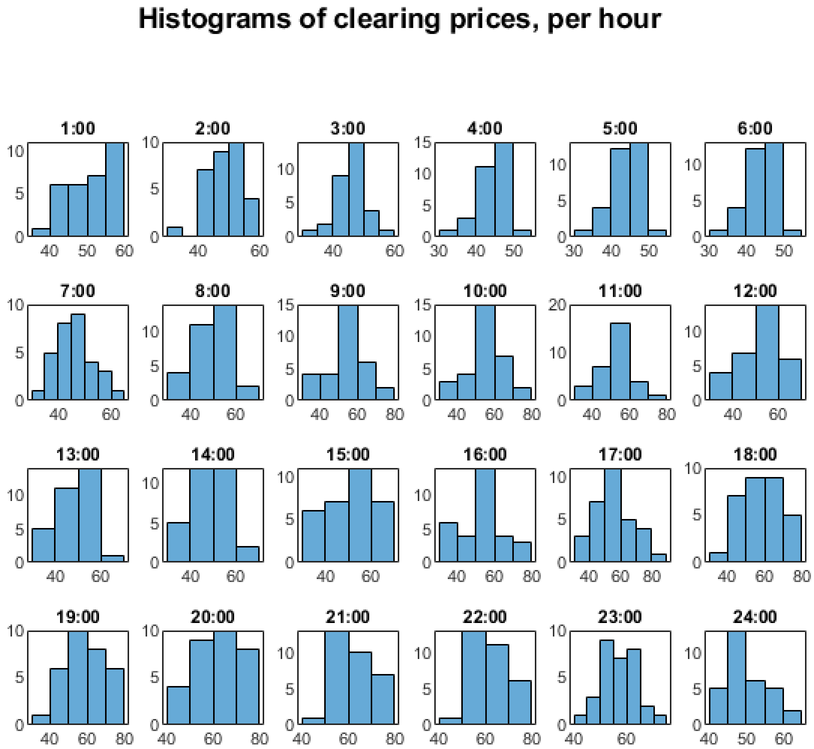


Figure 5: Histogram of clearing prices, separated by hour. **Source:** GME.

6.3 Market clearing algorithm

The detailed data from the GME allows the researcher to replicate the market clearing algorithm with accuracy. Most market operators do not publicly provide as much information as the Italian authority, so calculating precisely the equilibrium prices as an external researcher becomes either a very cumbersome task or even impossible[12]. The market authority solves a constrained welfare maximization problem whose gist may be written as:

$$\max_{\text{Prices, Quantities, Agents}} Welfare$$

$$s.t. \begin{cases} \text{Physical constraints} \\ \text{Merit-order constraints} \\ \text{Equilibrium constraints} \end{cases}$$

It is easier to develop the intuition from the simple case where there is only one market and no constraints. Welfare maximization here consists of crossing the aggregate supply and the aggregate demand functions to obtain a pair of equilibrium prices and quantities. It is known that this point maximizes the sum of the consumers' and producers' welfare.

The actual problem is more complicated because the authority deals with multiple zones (each region of the country) and constraints on transmission capacity and individual priorities. It must now find a unique price that maximizes the total sum of consumer and producer welfare. I refer the reader to Savelli et al (2017) and Savelli et al (2018) for a full description of the market clearing algorithm for the Italian market and its complete derivation.

In practice, the GME solves this by iterating two steps until convergence: first, it solves the zonal problems and saves the Karush–Kuhn–Tucker multipliers; in turn, these are used to solve the problem of selecting a unique price for the whole Italian market, that feed the first step. A complete description of the official algorithm used by the GME can be found in the documentation of the EUPHEMIA algorithm[13]. Savelli et al (2017) and Savelli et al (2018) show that one can cast this two-step, iterative optimization problem as a single, mixed-integer linear programming problem. This is convenient because off-the-shelf optimization softwares can be used to solve it. Figure 6 shows the algorithm indeed recovers the prices provided by the GME.

[12] Reguant (2014) offers an approximation to the Spanish market clearing algorithm that tracks the published prices closely, at the expense of large computational costs. Other market operators (like the California ISO, the Northeast ISO, or the New York ISO) do not provide regional data, and ignoring this dimension leads to inaccurate computations of equilibrium prices and quantities. In these situations, one can use a functional estimation procedure as developed in Benatia (2018).

[13] Available at: https://www.mercatoelettrico.org/En/MenuBiblioteca/Documenti/20131108EuphemiaNov2013.pdf

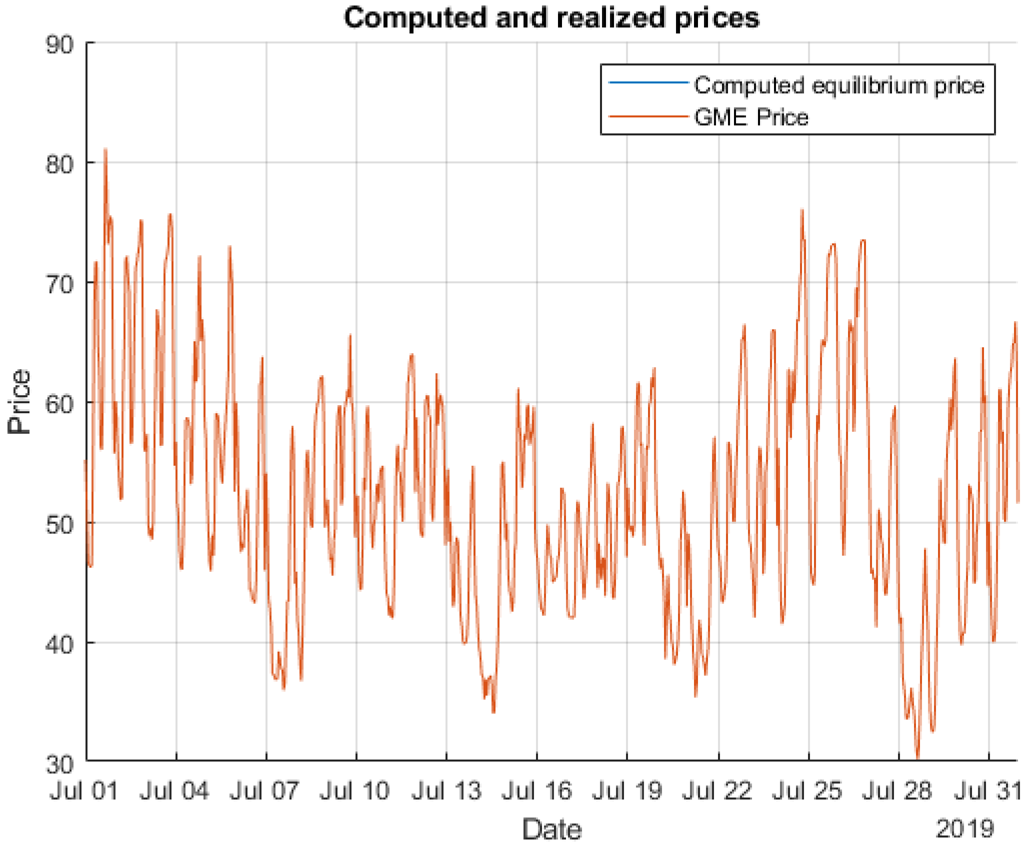


Figure 6: Time series of computed and actual prices. **Source:** GME and own computations.

6.4 Estimation

Recall that estimation of the model involves the following first-order conditions:

$$\nu(y(p,s_i),s_i) = p - y(p,s_i)\frac{\frac{\partial H}{\partial y}(p,y(p,s_i))}{\frac{\partial H}{\partial p}(p,y(p,s_i))}$$

For the demand side, and

$$\xi(x(p,\psi_j),\psi_j) = p + x(p,\psi_j)\frac{\frac{\partial \tilde{H}}{\partial x}(p,x(p,\psi_j))}{\frac{\partial \tilde{H}}{\partial p}(p,x(p,\psi_j))}$$

for the supply side.

It is worth pointing out that, even though the calculation of the equilibrium price here seems much more complicated than the one presented in Section 2, they are equivalent from the point of view of the consumers and the suppliers. Ultimately, the market clearing algorithm takes all bids and asks, and picks those that satisfy the solution. This creates a residual supply (from the point of view of the a buyer) or a residual demand (from the point of view of a seller) which is the object of interest, whose intersection with their submitted schedules results in the equilibrium price. Therefore, the main objects involved are the probability distributions H and $\tilde{H}$ exactly as defined in Section 2.

Calculating the market clearing prices accurately means that one can effectively estimate these using the bootstrap algorithms in Section 4, which now must include one extra step:

Algorithm 5. *(Modified bootstrap algorithm for the buyers)*

1. *Fix buyer i.*
2. *Repeat the following steps B times:*
   - *Resample N suppliers' and $M-1$ clients' schedules, excluding i.*
   - *Solve the market clearing algorithm with the re-sampled supply and demand functions.*
   - *Compute and store the residual supply, RS^b_{-i}.*
3. *Compute the partial derivatives H_p and H_y numerically, obtaining $\hat{H}_p$ and $\hat{H}_y$. This is done with the formulas from Section 3.*
4. *Compute the estimated marginal value at each price p_k:*

$$\hat{\nu}(p_k, y_i(p_k)) = p_k - y_i(p_k)\frac{\hat{H}_y}{\hat{H}_p}$$

Algorithm 6. *(Modified bootstrap algorithm for the sellers)*

1. *Fix seller j.*
2. *Repeat the following steps B times:*
   - *Resample M clients' $N-1$ suppliers' schedules, excluding j.*
   - *Solve the market clearing algorithm with the re-sampled supply and demand functions.*
   - *Compute and store the residual demand, RD^b_{-j}.*
3. *Compute the partial derivatives $\tilde{H}_p$ and $\tilde{H}_y$ numerically, obtaining $\hat{\tilde{H}}_p$ and $\hat{\tilde{H}}_y$. This is done with the formulas from Section 3.*
4. *Compute the estimated marginal value at each price p_k:*

$$\hat{\xi}(p_k, x_j(p_k)) = p_k + x_j(p_k)\frac{\hat{\tilde{H}}_x}{\hat{\tilde{H}}_p}$$

Solving the GME problem in each bootstrap repetition is necessary because different bids and ask functions may make different restrictions bind; consequently, accurate construction of the residual supply and the residual demand curves requires it. Apart from that, the algorithms remain the same. Figures 7 and 8 show the estimation results for two individuals, one on the demand side and one on the supply side, in the 12-1 a.m. market. It's worthwhile to notice that both curves are to the right of their submitted schedules, which is interpreted as both of these agents exerting some market power.

Due to the lack of structure in my model, the estimated marginal value for the supplier is bumpy. Intuitively, it should also be monotonic in the offered quantity. Exploring the bootstrap simulations indicate that this is due to not including enough data in the estimation. Due to computational costs of including more observations, one possible avenue is to impose more structure in the estimation. For example, this can be done by applying

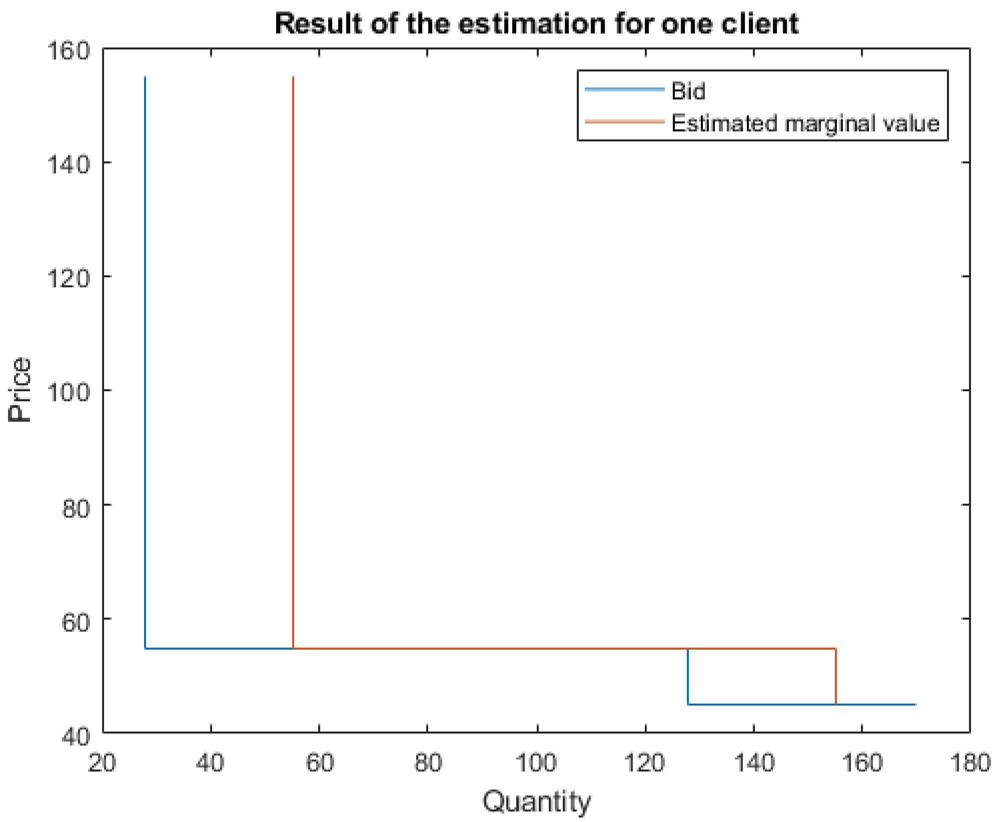


Figure 7: Result of the estimation for one client. **Source:** GME and own calculations.

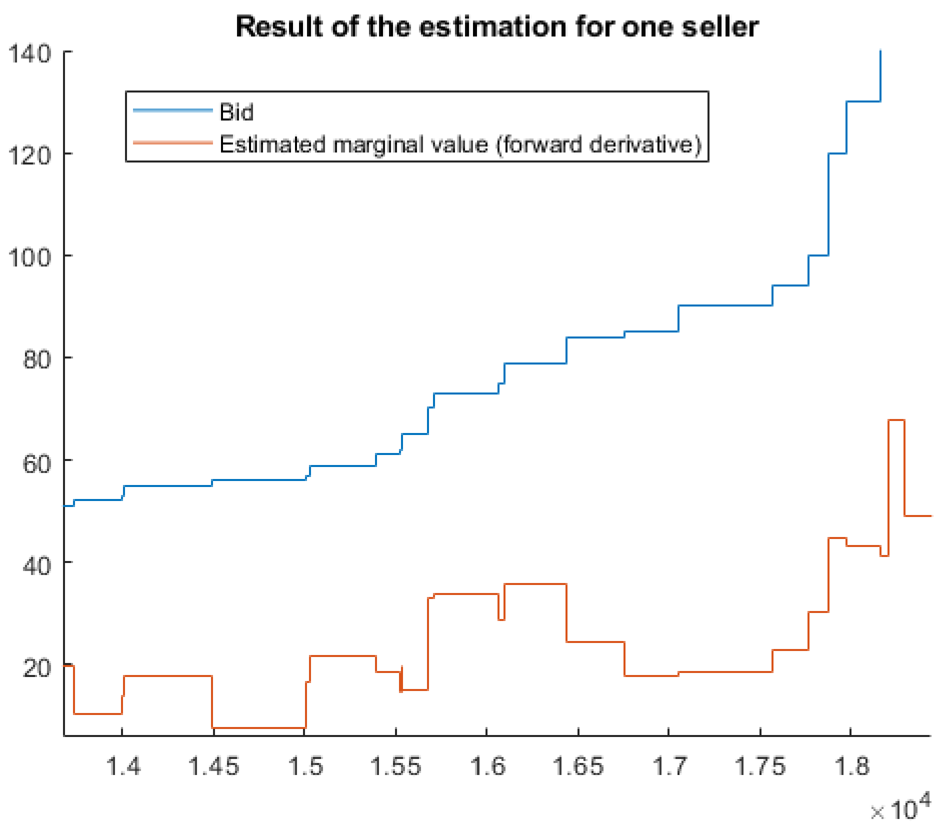


Figure 8: Result of the estimation for one seller. **Source:** GME and own calculations.

the estimation strategy presented in Matzkin's (2003) paper, as discussed previously in Section 3. In this case, one should notice that the first-order conditions can be written in the following form:

$$\nu_i(p, y_i) = m_c(p, s_i)$$
$$\xi_j(p, x_j) = m_s(p, \psi_j)$$

Under the assumption that m_c and m_s are monotonic in s_i and ψ_j, respectively; and a normalization on the distributions of the private signals, Matzkin's (2003) is applicable. Estimating the model this way is a future step in this work.

7 Conclusion

The estimation of marginal values is key to studying auctions. In this paper, I present a methodology to back out estimates for those in a two-sided multi-unit auction setup jointly and without imposing any functional form for the cost or demand functions. This methodology consists of a bootstrap algorithm to estimate a first-order condition as in Guerre, Perrigne, and Vuong (2000). In order to do so, the econometrician must have access to the individual schedules from both the demand and the supply sides and know how to calculate the market clearing prices.

If the bids are monotonic in price, the functions of interest are identified and the algorithms described in Section 4 may be used to estimate the marginal valuations. The Monte Carlo simulations show numerically that the procedure indeed generates estimates close to the true marginal values.

In the empirical application, I show that two agents (one client and one supplier) in the Italian electricity market exert market power by understating their willingness to buy or sell energy in order to manipulate the market price in their favor.

Other possible applications of this methodology include the study of after-hour stock markets and commodity markets organized in this anonymous, two-sided setup. For example, the NYSE conducts two-sided auctions this way after the typical trading hours end, and the price formed overnight is the opening price on the following day. A researcher interested in this would need to take care of the inter-temporal component of the decision, as one carries over their stocks. A starting point for this is Hortaçsu's (2002) study of Turkish bond auctions, which face the same problem.

7.1 Further work

There are many avenues to improve the work I did so far, as I pointed out during the exposition.

When it comes to the estimation procedure, I pointed out that it is possible to impose extra restrictions (monotonicity of the marginal valuation functions and a restriction in the shape of the private signals) to apply Matzkin's (2003) methodology. Imposing this extra structure allows the extraction of more information from the data, which I expect to result in estimatimated marginal valuation functions that are not as bumpy as the one I showed in Figure 7. Still on the topic of estimation, another approach is to apply a kernel estimator instead of resorting to the proposed resampling procedure. One expected advantage of said method is that it does not need to deal with the step-function nature of the data, which will possibly enhance my estimates as well. The Monte Carlo simulator I have set up is a great laboratory to test out and compare these different estimation methods.

There's no shortage of future work on the application side of this paper as well. With the estimated marginal valuations of both sides of the market, I can create a synthetic competitive benchmark where no agent exerts market power (or only suppliers, or only buyers for that matter). Since the total economic welfare of the agents is the value of the optimization problem described in Section 5, comparing different competitive scenarios and other counterfactuals are within reach.

Finally, one last point worth addressing in future work is the fact that the private signal may be correlated across hours for the same individual. Taking this into consideration in the estimation may be (potentially) complicated, but there is one structure that is favored by the data. As I pointed out in Section 5, all agents must submit their schedules for all the 24 hours prior to the market clearing. In this case, one can treat it as if the individual receives only one signal for the whole day, then they would solve a problem that resembles:

$$\max_{\{y_i^h(\cdot)\}_{i=1}^{24}} \mathbb{E}U_i(s_i) = \sum_{h=1}^{24} \int_0^{\infty} \left[\int_0^{y_i^h(p)} \nu_i(u)du - py_i^h(p) \right] dH_h(p, y_i^h(p))$$

Where $y_i^h(\cdot)$ is the submitted schedule for hour h of the day. Meaning the consumer submits all her schedules at once, taking into consideration the whole day. Estimating this is key to address this signal correlation problem, and its procedure has yet to be developed.

References

[1] David Benatia. Functional Econometrics of Multi-Unit Auctions: An Application to the New York Electricity Market. *Université de Montréal*.

[2] Simona Bigerna and Carlo Andrea Bollino. Demand market power and renewables in the Italian electricity market. *Renewable and Sustainable Energy Reviews*, 55:1154 – 1162, 2016.

[3] Natalia Fabra and Mar Reguant. Pass-Through of Emissions Costs in Electricity Markets. *American Economic Review*, 104(9):2872–2899, September 2014.

[4] Ali Hortaçsu and David McAdams. Empirical Work on Auctions of Multiple Objects. *Journal of Economic Literature*, 56(1):157–184, March 2018.

[5] Jakub Kastl. Discrete Bids and Empirical Inference in Divisible Good Auctions. *The Review of Economic Studies*, 78(3):974–1014, 2011. Publisher: [Oxford University Press, Review of Economic Studies, Ltd.].

[6] Rosa L. Matzkin. Nonparametric Estimation of Nonadditive Random Functions. *Econometrica*, 71(5):1339–1375, 2003. Publisher: [Wiley, Econometric Society].

[7] David McAdams. Partial identification and testable restrictions in multi-unit auctions. *Journal of Econometrics*, 146(1):74–85, September 2008.

[8] Mar Reguant. Complementary Bidding Mechanisms and Startup Costs in Electricity Markets. *The Review of Economic Studies*, 81(4 (289)):1708–1742, 2014. Publisher: [Oxford University Press, The Review of Economic Studies, Ltd.].

[9] Robert Wilson. Auctions of Shares. *The Quarterly Journal of Economics*, 93(4):675–689, 1979. Publisher: Oxford University Press.

[10] Frank A. Wolak. Identification and Estimation of Cost Functions Using Observed Bid Data: An Application to Electricity Markets. In Lars Peter Hansen, Mathias Dewatripont, and Stephen J. Turnovsky, editors, *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*, volume 2 of *Econometric Society Monographs*, pages 133–169. Cambridge University Press, Cambridge, 2003.

[11] Frank A. Wolak. Measuring Unilateral Market Power in Wholesale Electricity Markets: The California Market, 1998-2000. *American Economic Review*, 93(2):425–430, May 2003.

[12] Catherine D. Wolfram. Measuring Duopoly Power in the British Electricity Spot Market. *American Economic Review*, 89(4):805–826, September 1999.

[13] Iacopo Savelli, Bertrand Cornélusse, Antonio Giannitrapani, Simone Paoletti, and Antonio Vicino. A new approach to electricity market clearing with uniform purchase price and curtailable block orders. *Applied Energy*, 226:618–630, September 2018.

[14] Iacopo Savelli, Antonio Giannitrapani, Simone Paoletti, and Antonio Vicino. An Optimization Model for the Electricity Market Clearing Problem With Uniform Purchase Price and Zonal Selling Prices. *IEEE Transactions on Power Systems*, 33(3):2864–2873, May 2018. Conference Name: IEEE Transactions on Power Systems.

[15] David McAdams. Monotone Equilibrium in Multi-Unit Auctions. *The Review of Economic Studies*, 73(4):1039–1056, 2006. Publisher: [Oxford University Press, Review of Economic Studies, Ltd.].

[16] Devnath Shah and Saibal Chatterjee. A comprehensive review on day-ahead electricity market and important features of world's major electric power exchanges. 2020.

[17] Ali Hortaçsu and David McAdams. Mechanism Choice and Strategic Bidding in Divisible Good Auctions: An Empirical Analysis of the Turkish Treasury Auction Market. *Journal of Political Economy*, 118(5):833–865, October 2010. Publisher: The University of Chicago Press.

[18] Morton I. Kamien and Nancy L. Schwartz. *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management.* Number 31 in Advanced Textbooks in Economics.

[19] Bernard Elyakime, Jean Jacques Laffont, Patrice Loisel, and Quang Vuong. First-Price Sealed-Bid Auctions with Secret Reservation Prices. *Annales d'Économie et de Statistique*, (34):115–141, 1994. Publisher: [GENES, ADRES].

Appendix

Obtaining the first-order conditions

In this section, I show how to obtain the first order condition for both sides of the market.

Recall that we're working under the assumption that private signals are independent. To simplify the notation, let $y_i(p) \equiv y(p, s_i)$.

Demand side

Each buyer wants to solve the following problem:

$$\mathbb{E}U_i(s_i) = \int_0^\infty \left[\int_0^{y_i(p^c)} \nu_i(u)du - p^c y_i(p^c) \right] dG(p^c|y_i(p))$$

Where $G(p^c|y_i(p))$ is the distribution of equilibrium prices given the buyer's submitted demand curve. The dependence on their schedule comes from the fact that the buyers can influence the equilibrium price when, say, they submit a higher demand curve. One can also rewrite G in terms of the residual demand the individual faces in order to highlight this dependence and obtain the following probability distribution:

$$H(p, y_i(p)) = \mathbb{P}[y_i(p) \leq RS_{-i}(p)]$$

Where RS_{-i} is the residual supply faced by individual i, that is,

$$RS_{-i}(p) = \sum_{j=1}^{N} x_j(p) - \sum_{k=1,k\neq i}^{M} y_k(p)$$

From now on, I'll work with H instead of G. I follow the steps in Hortacsu (2002) to obtain the first-order condition. First, define: $\pi(p^c) = \pi(y_i(p^c)) \equiv \int_0^{y_i(p^c)} \nu_i(u)du - p^c y_i(p^c)$.

$$\frac{\partial \pi}{\partial p} = y'(p)\nu_i(y(p)) - py'(p) - y(p)$$

Now, we must have that $\pi(\infty) = 0$ because if price goes to infinity, the demand is zero. Under this, go back to the first equation and use integration by parts to write the expected utility as:

$$\mathbb{E}U_i = -\int_0^\infty (y'(p^c)\nu_i(y(p^c)) - p^c y'(p^c) - y(p^c))H(p^c, y(p^c))dp^c$$

It's time to take the derivatives to obtain the first order condition. This is done with Kamien and Schwartz's (1993) result:

$$F_y = \frac{d}{dp}F_{y'}$$

Where $F_y = (y'(p^c)\nu_i(y(p^c)) - p^c y'(p^c) - y(p^c))H(p^c, y(p^c))$

$$F_y = (1 - y'\nu')H + (-y'(\nu - p) + y)H_y$$

$$F_{y'} = -(\nu - p)H$$

$$\frac{d}{dp}F_{y'} = -(\nu'y' - 1)H - (\nu - p)[H_p + H_y y']$$

Equalize both to get the final version of the first-order condition,

$$yH_y = (-\nu H_p + PH_p)$$

$$yH_y = -(\nu - p)H_p$$

$$-yH_y = (\nu - p)H_p$$

$$\nu = p - y\frac{H_y}{H_p}$$

Supply side

Since the problem is symmetrical, obtaining the first-order condition for each agent on the supply side will follow the exact same steps. Start with the problem each of those agents want to solve:

$$\mathbb{E}\Pi_j(\psi_j) = \int_0^\infty \left[p^c x_j(p^c) - \int_0^{x_j(p^c)} \xi_j(u)du\right] dG(p^c|x_j(p^c))$$

Again, G is the probability distribution of the equilibrium prices. The equivalent probability distribution H is defined according to the residual demand supplier j faces:

$$H(p, x_j(p)) = \mathbb{P}[x_j(p) \leq RD_{-j}(p)]$$

Where $RD_{-j}(p)$ is given by:

$$RD_{-j}(p) = \sum_{i=1}^{M} y_i(p) - \sum_{k=1,k\neq j}^{N} x_k(p)$$

The goal is to rewrite the expected profit function by integrating by parts and only then take the first-order conditions. Define: $\pi(p^c) = \pi(x_j(p^c)) \equiv p^c x_j(p^c) - \int_0^{x_j(p^c)} \xi_j(u)du$.

$$\frac{\partial \pi}{\partial p} = px' + x - x'\xi_i(x)$$

Now, we must have that $\pi(0) = 0$, a direct implication that if price goes to zero, then the supply is zero. Under this, go back to the first equation and use integration by parts to get:

$$\mathbb{E}\Pi_j = -\int_0^\infty (p(p^c)x'(p^c) + x(p^c) - x'(p^c)\xi_j(x(p^c)))H(p^c, x_j(p^c))dp^c$$

Let $F \equiv -(px' + x - x'\xi_j(x))$ and use Kamien and Schwartz's (1993) result:

$$F_x = \frac{d}{dp}F_{x'}$$

$$F_x = (x'\xi - 1)H + (x + x'(p - \xi))H_x$$

$$F_{x'} = -(p - \xi)H$$

$$\frac{d}{dp}F_{x'} = -(1 - \xi'x')H - (p - \xi)[H_p + H_x x']$$

Equalize both to get the final version of the first order condition,

$$(1 - \xi'x')H + (p - \xi)[H_p + H_x x'] = -(x'\xi - 1)H - (x + x'(p - \xi))H_x$$

$$\xi = p + x\frac{H_x}{H_p}$$

A (surprisingly nice) solution

To provide proper Monte Carlo simulations, I obtain a solution to the model. Typically, this is complicated because the probability distributions H are hard objects to deal with. Of course one could obtain these solutions by using numerical methods, but I show that if the private values follow an exponential distribution, this problem has a linear solution under a restriction on the parameters of the model.

The proof strategy is guess-and-verify[14]. Assume that the true demand function of a buyer is:

$$D(p, s_i) = \alpha + \beta s_i + \gamma p,\ \alpha > 0, \gamma < 0$$

The inverse demand gives us the true valuation of the individual:

$$\nu(q, s_i) = \frac{1}{\gamma}(q - \alpha - \beta s_i)$$

On the supply side, things work similarly. The true supply function is:

$$S(p, \psi_j) = \tilde{\alpha} + \tilde{\beta}\psi_j + \tilde{\gamma} p,\ \tilde{\alpha} > 0, \tilde{\gamma} > 0$$

The inverse supply gives the true valuation of the supplier:

$$\xi(q, \psi_j) = \frac{1}{\tilde{\gamma}}(q - \tilde{\alpha} - \tilde{\beta}\psi_j)$$

The guesses are that the bid and ask functions, determined by the true parameters in the model, are linear in the private signals and prices. That is,

$$y(p, s_i) = a + bs_i + cp$$

$$x(p, \psi_j) = \tilde{a} + \tilde{b}\psi_j + \tilde{c}p$$

I will show that this is indeed a solution if the private values are i.i.d. exponential and we impose that $b = -\tilde{b}$.

Preliminaries

Let the private values be distributed as:

$$s_i \sim Exponential(\lambda)$$

$$\psi_j \sim Exponential(\lambda)$$

This is convenient because the distribution of the random variable

$$\sum_{j=1}^{N} \psi_j + \sum_{k=1, k\neq i}^{M} s_i$$

Follows an Erlang distribution with parameters $(\lambda, N + M - 1)$.[15]

Similarly,

$$\sum_{i=1}^{M} s_i + \sum_{k=1, l\neq j}^{N} \psi_k \sim Erlang(\lambda, M + N - 1)$$

This will become very convenient because the probability density function of a variable $T \sim Erlang(\lambda, k)$ is given by:

$$f(t; \lambda, k) = \frac{\lambda^k t^{k-1} e^{-\lambda t}}{(k-1)!}$$

[14] I follow Hortacsu's (2001) approach to obtain this solution.

[15] See Forbes et al (2011.)

Demand side

I use a guess-and-verify strategy to obtain a solution. The guess is that $y_i(p) = a + bs_i + cp$ and $x_j(p) = \tilde{a} + \tilde{b}\psi_j + \tilde{c}p$. Recall that the central object of the first-order condition of an individual on the demand side is the probability distribution H, so use the guess in its definition:

$$\mathbb{P}(y_i(p) \leq RS_{-i}(p)) = \mathbb{P}\left(y_i(p) \leq \sum_j x_j(p) - \sum_{k \neq i} y_k(p)\right)$$

$$\mathbb{P}(y_i(p) \leq RS_{-i}(p)) = \mathbb{P}\left(y_i(p) \leq N\tilde{a} + \tilde{b}\sum_j \psi_j + \tilde{c}Np - (M-1)a - b\sum_{k \neq i} s_k - c(M-1)p\right)$$

Now, if the parameters of the model are such that $b = -\tilde{b}$, we can solve for the Erlang random variable:

$$\mathbb{P}(y_i(p) \leq RS_{-i}(p)) = \mathbb{P}\left(\frac{y_i(p) - N\tilde{a} + (M-1)a - \tilde{c}Np + c(M-1)p}{-b} \leq \sum_j \psi_j + \sum_{k \neq i} s_k\right)$$

$$\mathbb{P}(y_i(p) \leq RS_{-i}(p)) = 1 - \mathbb{P}\left(\frac{y_i(p) - N\tilde{a} + (M-1)a - \tilde{c}Np + c(M-1)p}{-b} \geq \sum_j \psi_j + \sum_{k \neq i} s_k\right)$$

From the previous subsection, we know that this is an Erlang distribution with parameters $(\lambda, M+N-1)$. To simplify the notation, let $t \equiv \frac{y_i(p) - N\tilde{a} + (M-1)a - \tilde{c}Np + c(M-1)p}{-b}$. Therefore,

$$\mathbb{P}(y_i(p) \leq RS_{-i}(p)) = \sum_{n=0}^{M+N-2} \frac{e^{-\lambda t}(\lambda t)^n}{n!}$$

Taking the partial derivatives,

$$H_y = \frac{\lambda^{M+N-1} t^{M+N-2} e^{-\lambda t}}{(M+N-2)!}\left(-\frac{1}{b}\right)$$

$$H_p = \frac{\lambda^{M+N-1} t^{M+N-2} e^{-\lambda t}}{(M+N-2)!}\left(-\frac{c(M-1) - \tilde{c}N}{b}\right)$$

Plug this back into the first-order condition to obtain:

$$\nu = p - y\frac{H_y}{H_p}$$

$$\nu = p - y\frac{\frac{\lambda^{M+N-1} t^{M+N-2} e^{-\lambda t}}{(M+N-2)!}\left(-\frac{1}{b}\right)}{\frac{\lambda^{M+N-1} t^{M+N-2} e^{-\lambda t}}{(M+N-2)!}\left(-\frac{c(M-1) - \tilde{c}N}{b}\right)}$$

$$\nu = p - \frac{y}{c(M-1) - \tilde{c}N}$$

For ease of notation, define $\varphi^{-1} \equiv -[c(M-1) - \tilde{c}N]$. So that $\nu = p + \varphi y$.

So let's now work with the guesses:

$$a + bs_i + cp = \alpha + \beta s_i + \gamma(p + \varphi y)$$

$$a + bs_i + cp = \alpha + \beta s_i + \gamma p + \gamma\varphi y$$

Plug in again the guess for y,

$$a + bs_i + cp = \alpha + \beta s_i + \gamma p + \gamma\varphi[a + bs_i + cp]$$

$$a + bs_i + cp = [\alpha + \gamma\varphi a] + [\beta + \gamma\varphi b]s_i + [\gamma + \gamma\varphi c]p$$

Comparing coefficients,

$$\begin{cases} a = \alpha + \gamma\varphi a \Rightarrow a = \frac{\alpha}{1-\gamma\varphi} \\ b = \beta + \gamma\varphi b \Rightarrow b = \frac{\beta}{1-\gamma\varphi} \\ c = \gamma + \gamma\varphi c \Rightarrow c = \frac{\gamma}{1-\gamma\varphi} \end{cases}$$

Supply side

Let's now work with the supply side. The steps, however, are the same.

$$\mathbb{P}(x_j(p) \leq RD_{-j}(p)) = \mathbb{P}\left(x_j(p) \leq \sum_{i=1}^{M} y_i(p) - \sum_{k=1,k\neq j}^{N} x_k(p)\right)$$

$$\mathbb{P}(x_j(p) \leq RD_{-j}(p)) = \mathbb{P}\left(x_j(p) \leq -(N-1)\tilde{a} - \tilde{b}\sum_{k=1,k\neq j}^{N}\psi_k - \tilde{c}(N-1)p + Ma + b\sum_{i=1}^{M} s_i + cMp\right)$$

With $\tilde{b} = -b$, and solving for the RV

$$\mathbb{P}(x_j(p) \leq RD_{-j}(p)) = \mathbb{P}\left(\frac{x_j(p) + (N-1)\tilde{a} + \tilde{c}(N-1)p - Ma - cMp}{b} \leq \sum_{k=1,k\neq j}^{N}\psi_k + \sum_{i=1}^{M} s_i\right)$$

$$\mathbb{P}(x_j(p) \leq RD_{-j}(p)) = 1 - \mathbb{P}\left(\frac{x_j(p) + (N-1)\tilde{a} + \tilde{c}(N-1)p - Ma - cMp}{b} \geq \sum_{k=1,k\neq j}^{N}\psi_k + \sum_{i=1}^{M} s_i\right)$$

Again, let $t \equiv \frac{x_j(p)+(N-1)\tilde{a}+\tilde{c}(N-1)p-Ma-cMp}{b}$. One can rewrite the probability distribution of interest as:

$$\mathbb{P}(x_j(p) \leq RD_{-j}(p)) = \sum_{n=0}^{M+N-2} \frac{e^{-\lambda t}(\lambda t)^n}{n!}$$

The ratio of the partial derivatives is:

$$\frac{H_x}{H_p} = \frac{1}{\tilde{c}(N-1) - cM}$$

Plugging this into the first-order condition of the j^{th} supplier,

$$\xi = p + x\frac{H_x}{H_p} = p + \frac{x}{\tilde{c}(N-1) - cM}$$

Define $\tilde{\varphi}^{-1} \equiv \tilde{c}(N-1) - cM$. Then $\xi = p + \tilde{\varphi}x$. Let's go back to the guesses:

$$\tilde{a} + \tilde{b}\psi_j + \tilde{c}p = \tilde{\alpha} + \tilde{\beta}\psi_j + \gamma(p + \tilde{\varphi}x)$$

$$\tilde{a} + \tilde{b}\psi_j + \tilde{c}p = \tilde{\alpha} + \tilde{\beta}\psi_j + \tilde{\gamma}p + \tilde{\gamma}\tilde{\varphi}x$$

Plug in again the guess for x,

$$\tilde{a} + \tilde{b}\psi_j + \tilde{c}p = \tilde{\alpha} + \tilde{\beta}\psi_j + \tilde{\gamma}p + \tilde{\gamma}\tilde{\varphi}[\tilde{a} + \tilde{b}\psi_j + \tilde{c}p]$$

$$\tilde{a} + \tilde{b}\psi_j + \tilde{c}p = [\tilde{\alpha} + \tilde{\gamma}\tilde{\varphi}\tilde{a}] + [\tilde{\beta} + \tilde{\gamma}\tilde{\varphi}\tilde{b}]\psi_j + [\tilde{\gamma} + \tilde{\gamma}\tilde{\varphi}\tilde{c}]p$$

Comparing coefficients,

$$\begin{cases} \tilde{a} = \tilde{\alpha} + \tilde{\gamma}\tilde{\varphi}\tilde{a} \Rightarrow \tilde{a} = \frac{\tilde{\alpha}}{1-\tilde{\gamma}\tilde{\varphi}} \\ \tilde{b} = \tilde{\beta} + \tilde{\gamma}\tilde{\varphi}\tilde{b} \Rightarrow \tilde{b} = \frac{\tilde{\beta}}{1-\tilde{\gamma}\tilde{\varphi}} \\ \tilde{c} = \tilde{\gamma} + \tilde{\gamma}\tilde{\varphi}\tilde{c} \Rightarrow \tilde{c} = \frac{\tilde{\gamma}}{1-\tilde{\gamma}\tilde{\varphi}} \end{cases}$$

Therefore, the solution of the model is linear in the private values and the price.

Solving the non-linear system

The definitions of φ and $\tilde{\varphi}$ imply that the non-linear systems in the two previous subsections are connected. Its solution maps a set of structural parameters $(\alpha, \beta, \gamma, \tilde{\alpha}, \tilde{\beta}, \tilde{\gamma}, M, N)$ onto optimal bid and ask functions.

$$\begin{cases} \tilde{a} = \frac{\tilde{\alpha}}{1-\tilde{\gamma}\tilde{\varphi}} \\ \tilde{b} = \frac{\tilde{\beta}}{1-\tilde{\gamma}\tilde{\varphi}} \\ \tilde{c} = \frac{\tilde{\gamma}}{1-\tilde{\gamma}\tilde{\varphi}} \\ a = \frac{\alpha}{1-\gamma\varphi} \\ b = \frac{\beta}{1-\gamma\varphi} \\ c = \frac{\gamma}{1-\gamma\varphi} \\ b = -\tilde{b} \\ \varphi^{-1} \equiv -[c(M-1) - \tilde{c}N] \\ \tilde{\varphi}^{-1} \equiv \tilde{c}(N-1) - cM \end{cases}$$

Which one can solve for the desired functions.

Hortaçsu's (2001, 2002) simpler setup

Hortacsu (2001, 2002) works with a simpler situation that is the closely related to the model I presented in Section 2. There are two simplications: the first is that there's only one supplier, who is willing to sell a specific quantity (normalized to a unit) at any price; the second is that this is a "pay-as-you-go" setup[16].

The first-order condition for a buyer is:

$$\nu(y(p, s_i), s_i) = p - \frac{\frac{\partial H}{\partial y}(p, y(p, s_i))}{\frac{\partial H}{\partial p}(p, y(p, s_i))}$$

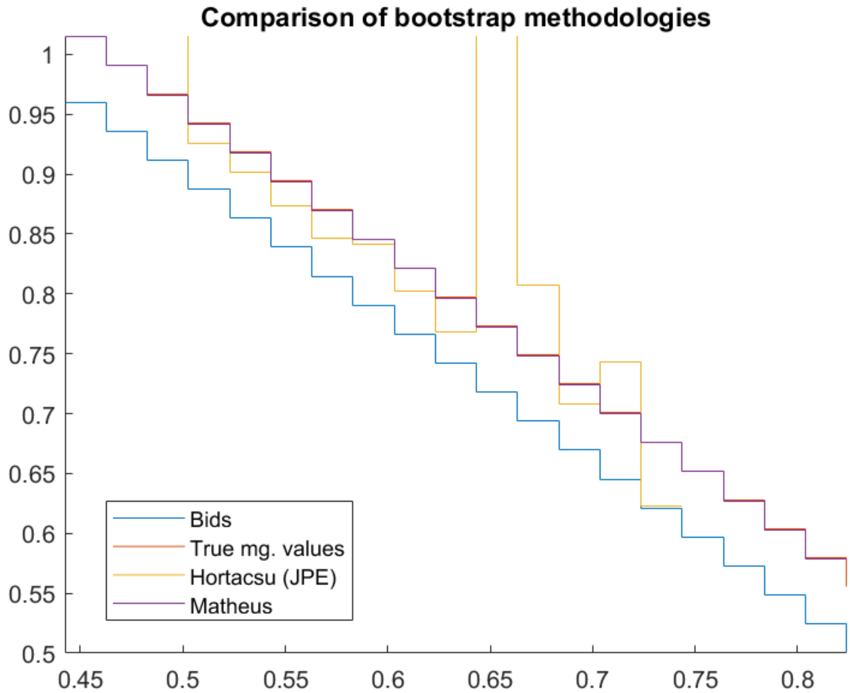
Which can be estimated with the procedures in this paper, using the same algorithms described in Section 4. In this situation, his methodology can easily be implemented by doing a kernel estimation of the equilibrium price distribution obtained from the bootstrap procedure.

This procedure, however, is unlikely to be accurate if the steps in the demand function are far from each other. The reason is that for each bootstrap repetition, the equilibrium quantity is not held fixed.

Ultimately, this means that the quantity he is effectively calculating is not a partial derivative, since the second argument of the H function is not held fixed.

Here, I simulate $1,000$ auctions with two buyers, $m = 2$. Following his paper, the distribution of private values is such that the demand curves are linear and the true marginal values are given by the demand curve added by a constant.

[16]See Wilson (1979).


Comparison of bootstrap methodologies
Bids
True mg. values
Hortacsu (JPE)
Matheus
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.5
0.55
0.6
0.65
0.7
0.75
0.8