Efficient Local IV Estimation of an Empirical Auction Model

HAN HONG and DENIS NEKIPELOV[1]

Key words: Semiparametric efficiency bound, local treatment effect, natural experiment, online auction, early jump bid

1 Introduction

Semiparametric efficiency and endogeneity are both important issues in the estimation of treatment effect models. Many papers in the literature, including Chernozhukov and Hansen (2005) and Newey (1990), have developed efficent estimators under conditional mean or median independence or quantile independence assumptions to account for endogenous regressors. Under the strong ignorability assumption, Hahn (1998) and Hirano, Imbens, and Ridder (2003) derived the semiparametric efficiency bound and developed semiparametric efficient estimators for the averaged treatment effect and the averaged treatment effect on the treated, while Firpo (2003) extended the analysis to quantile treatment effects.

In this paper we derive semiparametric efficiency bounds and efficient estimators for the conditional monotone local instrument variable (LIV) model studied in Abadie, Angrist, and Imbens (2002). We apply the semiparametrically efficient estimation method to analyze the relation between bid dispersion and early bidding in an online auction dataset, which is collected from a natural experiment conducted in Nekipelov (2007), The results confirm the theoretical findings developed in Nekipelov (2007). The semiparametric efficient estimation procedure substantially improves the statistical significance of the effect of jump bidding on bid dispersion.

The causal parameters identified under the LIV assumption, first advanced by Angrist (2004), are fundamentally different from those under conventional orthogonality assumption between the instrument and the structural errors. The validity of the LIV assumption

[1]Department of Economics, Stanford University and Department of Economics, Duke University, USA. The authors acknowledge generous research supports from the NSF, the Sloan Foundation and the Micro Incentives Research Center of Duke University. The usual disclaimer applies.

depends crucially on the context of the empirical application. With the exception of Frolich (2006), semiparametric efficiency implications of the LIV assumption have not been investigated. This paper takes a first step in developing an efficient estimator for a conditional linear model under the LIV assumption. Hong and Nekipelov (2007) extends these results to general nonlinear moment condition models for both unconditional and conditional parameters.

The baseline LIV model has a binary endogenous regressor and a binary instrument variable. The instrument variable weakly changes the endogenous regressor in one direction and identifies that the entire distributional causal effect for the *compiler* population where the endogenous regressor changes from zero to one as the instrumental variable changes from zero to one.

The efficient local IV estimator is suitable for analyzing the empirical implication of the model developed by Nekipelov (2007) on the relation between bid dispersion and early bidding in an online auction dataset. The theoretical model developed in Nekipelov (2007) showed that early bidding has competing effects on the dispersion of bids. On the one hand, early bidding deters entry and decreases the bid dispersion. On the other hand, early bidding provides more information and potentially increases both learning and bid dispersion. The incentive for early bidding is weaker when learning prevention dominates entry deterrence, in which case early jump should tend to increase bid dispersion. However, a simple regression of bid variation on an early bidding indicator suffers from endogeneity, because unobserved visibility of an auction item is correlated with both early jump bidding and bid dispersion. More visible items attract more determined well-informed bidders, who increase the price dispersion and reduce the incentive for early jump bids.

A local instrumental variable is given by an exogenous change of supply in a natural experiment conducted in Nekipelov (2007), who auctioned off additional supply of Robbie Williams' CD "The Greatest Hits" on ebay between October and December of 2006. Such exogenous increase in supply weakly increases the incentive for early jump bidding. We will therefore identify the set of *compliers* as the set of auctions which have no early jump bid prior to the supply increase but have early jump bid after the supply increases. By applying the efficient local IV estimator, we find that early jump bid has a significantly positive effect on the bid dispersion, implying that the incentive for learning prevention is stronger than the incentive for entry deterrence. Without controlling for endogeneity using the local instrument, early jump bid has a much smaller positive effect on bid dispersion. This confirms the presence of attenuation bias due to unobserved visibility which is negatively correlated with early jump bids but positively correlated with price dispersion.

In the rest of paper, section 2 describes the model and develops the efficiency bound and the efficient estimator, while section 3 applies the analysis to an auction model. Section 4 concludes.

2 Semiparametric efficiency bound in the linear model of treatment effects

2.1 Model

Consider the following model. There is a random vector $Y = (Y_1, Y_0)' \in \mathbb{R}^2$, a vector of binary variables $D = (D_1, D_0)' \in \{0, 1\} \times \{0, 1\}$, a binary instrument $Z \in \{0, 1\}$ and a vector of covariates $X \in \mathcal{X} \subset \mathbb{R}^k$. We use the assumptions of Abadie, Angrist, and Imbens (2002) to describe the distributions of the variables under consideration. Specifically, consider the following assumptions:

Assumption 1 *Almost everywhere in $\mathcal{X}$:*

- $(Y, D) \perp Z \,|\, X$
- $Pr(Z = 1 \mid X) \in (0, 1)$
- $E[D_1 \mid X] \neq E[D_0 \mid X]$
- $Pr(D_1 \geq D_0 \mid X) = 1$

The vector of variables (Y, D, Z, X) is only partially observable. We observe the dummy D_1 if $Z = 1$ and dummy D_0 if $Z = 0$. In this way a binary instrument can be interpreted as a random selection rule into a subgroup of observations where the probability of binary treatment is different from the main sample. Finally, we observe the outcome Y_1 if the observable dummy (either D_1 or D_0) is equal to 1 and Y_0 otherwise. Formally, we observe a vector (W, Z, X) whose values are determined by the values of the vector (Y, D, Z, X) as:

$$\begin{cases} w_1 = q_1\left(y, d, z, x\right) = y_1\, w_2 + y_0\, \left(1 - w_2\right) \\ w_2 = q_2\left(y, d, z, x\right) = d_0 + z\left(d_1 - d_0\right). \end{cases}$$

We are interested in estimation of the parameter vector $\beta \in \mathcal{B}$ (and especially its component β_0 reflecting the effect of binary dummy on the outcome variable) defining the moment function:

$$g(w, x, \beta) = w_1 - \beta_0 w_2 - x'\beta_1. \tag{1}$$

In addition, we are interested not in the effect for the entire population, but for a subpopulation for which selection Z unambiguously changes the treatment outcome, that is $D_1 > D_0$. In the treatment effect literature the members of such subpopulation are called "compliers". Equation (1) a conditional moment equation in the form:

$$\varphi\left(\beta, x, w_2\right) = E\left[g(w, x, \beta) \mid x,\, w_2,\, d_1 > d_0\right] = 0. \tag{2}$$

In a model with omitted variables, the treatment effect for compliers can be identified in the presence of an experiment which shifts the distribution of the binary treatment regressor but does not affect the distribution of omitted variable and the error term. Assume that the effect of the binary treatment regressor before and after the experiment is the same for compliers (observations shifted by the experiment) and can be described by a simple linear model:

$$Y_i = \beta_0 D_j + X'\beta_1 + \gamma\chi + \epsilon_i, \text{ for } i = 0,\, 1,\; j = 0,\, 1,$$

where χ is an omitted regressor and ϵ is the error term. The assumption that the model remains the same for observations affected by the experiment imposes the following moment restriction on the error term

$$E\left[\epsilon_i \,|\, x_i, w_{2i}, d_1 > d_2\right] = 0.$$

The binary regressor D is correlated with the omitted regressor χ. In general, omission of χ from the regression will produce a bias in the estimate of β_0. However, the structure of the experiment assures conditional independence between the omitted variable and the instrument

$$\chi \perp Z, \epsilon_i \,|\, X.$$

Abadie, Angrist, and Imbens (2002) demonstrated that for the subset of compliers Y is independent of W_2 given X: $Y \perp W_2 \,|\, X$. This is because for the subset of compliers: $Z = W_2$. This also implies that in the subset of compliers χ will be uncorrelated with W_2. Thus in the model

$$W_1 = \beta_0 W_2 + X'\beta + \gamma\chi + \eta, \quad \text{such that } D_1 > D_0$$

for the subset of compliers, the combined residual $\gamma\chi + \eta$ will be conditionally independent of W_2 given X. Therefore, this model will produce the conditional moment condition (2).

2.2 Semiparametric efficiency bound

We will assume that the outcome variable Y is continuous with conditional distributions of individual components for compliers $f\left(y_1 \mid d_1 > d_0, x\right)$ and $f\left(y_0 \mid d_1 > d_0, x\right)$. We denote the probabilities of binary variables by:

$$\begin{aligned}
&\mathcal{P}_0(x) = P\left(W_2 = 1 \mid Z = 0,\, X = x\right) = E\left[D_0 \mid x\right], \\
&\mathcal{P}_1(x) = P\left(W_2 = 1 \mid Z = 1,\, X = x\right) = E\left[D_1 \mid x\right], \\
&\mathcal{Q}(x) = E\left[Z \mid x\right], \quad \mathbf{P}\left(x\right) = P\left(X = x\right).
\end{aligned} \tag{3}$$

We can express the conditional probability of the binary treatment w_2 given the instrument in terms of the probabilities of treatment dummies as

$$P\left(W_2 = 1 \mid Z = z,\, X = x\right) = \mathcal{F}(z, x) = \mathcal{P}_1(x) z + \mathcal{P}_0(x)\left(1 - z\right).$$

The probability distribution of $W_2 \,|\, X$ without conditioning on Z can be expressed as an expectation over z given x so that $\mathcal{P}(x) = \mathcal{P}_1(x)\mathcal{Q}(x) + \mathcal{P}_0(x)\left(1 - \mathcal{Q}(x)\right)$. We will study the parameters of the moment condition (2) for the subset of compliers. Using the monotonicity assumption 1, the conditional density of the observable outcome W_1 for compliers can be expressed in terms of the observed conditional density of W_1 in the entire population and the probabilities listed in (3):

$$\begin{aligned} f_*\left(w_1 \mid x,\, w_2=1\right) &= f\left(y_1=w_1 \mid d_1>d_0,\, x\right) \\ &= \tfrac{\mathcal{P}_1(x)}{\mathcal{P}_1(x)-\mathcal{P}_0(x)} f\left(w_1 \mid w_2=1, z=1, x\right) - \tfrac{\mathcal{P}_0(x)}{\mathcal{P}_1(x)-\mathcal{P}_0(x)} f\left(w_1 \mid w_2=1, z=0, x\right), \end{aligned}$$

and

$$\begin{aligned} f_*\left(w_1 \mid x,\, w_2=0\right) &= f\left(y_0=w_1 \mid d_1>d_0, x\right) \\ &= \tfrac{1-\mathcal{P}_0(x)}{\mathcal{P}_1(x)-\mathcal{P}_0(x)} f\left(w_1 \mid w_2=0, z=0, x\right) - \tfrac{1-\mathcal{P}_1(x)}{\mathcal{P}_1(x)-\mathcal{P}_0(x)} f\left(w_1 \mid w_2=0, z=1, x\right). \end{aligned}$$

The conditional moment equation (2) can be expressed as an integral with the kernel function (1):

$$\begin{aligned} \varphi\left(\beta, x, w_2\right) &= \int g\left(w_1, w_2, x, \beta\right) f_*\left(w_1 \mid w_2,\, x\right)\, dw_2 \\ &= \int \left[w_1-\beta_0 w_2 - x'\beta\right] f_*\left(w_1 \mid w_2,\, x\right)\, dw_2. \end{aligned} \tag{4}$$

If X is a continuous random variable, this conditional moment equation is equivalent to the many unconditional moments for fixed values of x and w_2. Next, we will first find the semiparametric efficiency bound for this model.

Theorem 1 *Suppose that the selection probability functions in (3) have non-degenerate support on $\mathcal{X}$, and the model satisfies the standard regularity conditions. Denote $\omega_{w_2,z}(x) = V\left(w_1 \mid w_2, z, x\right)$ and $\gamma_{w_2,z}(x) = E\left(w_1 \mid w_2, z, x\right) - w_2\beta_0 - \beta_1' x$. Introduce matrix $\overline{\Omega}(x)$ with the following elements:*

$$\overline{\Omega}_{11}(x) = \left(\tfrac{\mathcal{P}_1(x)\omega_{11}(x)}{\mathcal{Q}(x)} + \tfrac{\mathcal{P}_0(x)\omega_{10(x)}}{1-\mathcal{Q}(x)} + \tfrac{\gamma_{11}^2(x)\mathcal{P}_1(x)\mathbf{P}(x)}{\mathcal{P}_0(x)\mathcal{Q}(x)(1-\mathcal{Q}(x))}\left[1-\tfrac{\mathcal{P}_1(x)\mathcal{P}_0(x)}{\mathbf{P}(x)}\right]\right),$$

$$\overline{\Omega}_{22}(x) = \left(\tfrac{(1-\mathcal{P}_1(x))\omega_{01}(x)}{\mathcal{Q}(x)} + \tfrac{(1-\mathcal{P}_0(x))\omega_{00}(x)}{1-\mathcal{Q}(x)} + \tfrac{\gamma_{00}^2(x)(1-\mathcal{P}_0(x))(1-\mathbf{P}(x))}{\mathcal{Q}(x)(1-\mathcal{Q}(x))(1-\mathcal{P}_1(x))}\left[1-\tfrac{(1-\mathcal{P}_0(x))(1-\mathcal{P}_1(x))}{1-\mathbf{P}(x)}\right]\right),$$

and

$$\overline{\Omega}_{21}(x) = \overline{\Omega}_{12}(x) = \left(\tfrac{\mathcal{P}_1(x)(1-\mathcal{P}_0(x))}{\mathcal{Q}(x)(1-\mathcal{Q}(x))}\gamma_{11}(x)\gamma_{00}(x)\right).$$

Then we can express the semiparametric efficiency bound for the model defined in (1) and (2) under assumption 1 as:

$$V_\beta = E\left\{(\mathcal{P}_1(x) - \mathcal{P}_0(x))^2 \begin{bmatrix} 1 & 0 \\ x & x \end{bmatrix} \overline{\Omega}(x)^{-1} \begin{bmatrix} 1 & x' \\ 0 & x' \end{bmatrix}\right\}^{-1}.$$

Moreover, this bound can be achieved in a GMM model for the subset of compliers for a moment function:

$$m(w, x, \beta) = (\mathcal{P}_1(x) - \mathcal{P}_0(x)) \begin{bmatrix} 1 & 0 \\ x & x \end{bmatrix} \overline{\Omega}^{-1}(x) \begin{pmatrix} \frac{w_2}{\mathbf{P}(x)} \\ \frac{1-w_2}{1-\mathbf{P}(x)} \end{pmatrix} \left[w_1 - \beta_0 w_2 - x'\beta_1\right].$$

The proof of this theorem is given in the appendix.

The main practical implication of Theorem 1 is that we can estimate the parameters efficiently in a relatively simple two step GMM framework. In the first step, we can use the basic result in Abadie, Angrist, and Imbens (2002) to find initial consistent estimate of coefficient β. Then in step two, we form an estimate of the weighting matrix, and find the efficient estimate of β.

In the next section we will apply the efficient GMM procedure to analyzing the relation between bid variation and early bidding on eBay auction. The structure of our data allows us to find a binary instrument, which satisfies the requirements of the monotone local instrumental variable model.

3 Early bidding and the variance of bids on eBay

3.1 Description of the problem

In Nekipelov (2007) the author develops a continuous-time theoretical model of bidding on eBay auctions that is aimed at explaining early and multiple bidding on eBay. The model is designed to capture two important characteristics of Internet auctions: multiplicity of listings of similar items, and uncertainty of bidders about the number and skills of their rivals. This situation is modeled by a single agent optimization problem in which the bidder maximizes her expected surplus from winning the auction by submitting bids.

The dynamics of the auction is characterized by a stochastic price process driven by Poisson jumps, such that the bidder under consideration can affect the size of the price jumps. The frequency of price jumps is related to the multiplicity of listings on eBay. In a more general setting of a two stage bidding model, at the first stage the bidder chooses the auction, and bids in the chosen auction at the second stage. Entry in the first stage will depend on the price in the auction. It is assumed that the frequency of price jumps depends on price, time and the latent *visibility* of the auction. The latent visibility of the auction determines the idiosyncratic differences between different auctions, and is assumed to be an unobserved fundamental characteristic of the auction[2]. If in the beginning of the auction bidders have different information about the visibility, the expected payoff of bidders with more precise information will have higher expected payoffs. For this reason, bidders will have an incentive to learn about the visibility during the auction.

The analysis of this model shows that there are two important features of bidding behavior in such auction. The first feature is connected with the fact that on eBay, the multiplicity of auction listings makes the entry of bidders into auctions endogenous. More specifically, in the presence of multiple simultaneous auctions the entry into auctions depends on prices. Thus, the incumbent bidders can prevent entry of other bidders by raising the price. In this case, bidding early will reduce the number of potential entrants, but will also decrease the expected surplus of the incumbent bidder from winning. The second feature is connected with the heterogeneity of bidders' experience. As bidders have imperfect information about their rivals, they will need to "experiment" with the response of the rivals on bidding. For this reason, more experienced bidders will tend to bid less frequently and close to the end of the auction to prevent less experienced bidders from learning about the rivals.

The variation of the visibility across auction is correlated with observable characteristics of the auction which are functions of the visibility, including price variance and frequency of jump bids. For this reason, regression estimation of relationships between observable variables across auctions will be prone to bias. The identification methodology which we have developed in section 2 can resolve this problem if we have an auxiliary dataset in which the observable instruments of the auction market are significantly changed while

[2]To achieve identification we assume that the frequency of price jumps is monotone increasing in the visibility of the auction, and that the visibility takes values in a closed (and known) segment of a real line.

the structure of the unobserved heterogeneity is the same as in the original sample. The presence of such dataset allows us to create an instrument and remove the endogeneity bias.

An interesting relationships which can be studied on eBay auctions is the connection between the variance of bids in the auction and the presence of an early jump bid. Under the private value assumption, in the case of sealed-bid auctions, early jump bids should not change the variance of bids in the auction. In the ascending price button auctions, early jump bids should reduce the variance of bids because they truncate the distribution of the bids. The pattern on eBay auctions is very different. More specifically, the presence of early bids can increase the variance of bidding. We note the following relation: both the variance of bids and the early jump bidding depend on the entry rate of the bidders. The entry rate, in turn, is determined by the unobserved visibility of the auction. The theoretical model in Nekipelov (2007) demonstrates that early jump bidding depends on the elasticity of the entry rate with respect to the price. This elasticity should be smaller for auctions with high visibilities (intuitively, this suggests that more visible auctions have a more persistent entry of bidders). Thus, we should expect that the probability of early jump bidding decreases with the visibility of the auction. The variance of bids is determined by the number of bidders, which is increasing on average as the visibility (and, thus, the entry rate) increases. For this reason, we should expect that the relation of both variables with unobserved visibility should induce an attenuation bias in the estimate of the coefficient for the dependence between the variance of bids and the early jump bid dummy.

To correct this bias we will employ the data from the field experiment. In the field experiment the market for pop-music CDs was inflated which changed the entry patterns by increasing the elasticity of the entry rate with respect to the price. If we use the data for two periods before and after the market expansion, we can use the dummy for the inflated market as a binary instrument in the estimation the relationship between the variance of bids and early jump bidding (While the field experiment has changed the bidder entry rate, it does not changed visibilities of auctions).

3.2 Data

In the design of the field experiment in Nekipelov (2007), three important goals were pursued. First, it was necessary to find an item whose valuations are uncorrelated across bidders, for which standard independent private-value auction is valid. Second, the market for item should be active so that enough sales can be recorded in a month-long period. Third, the marked can not be too large as inflating a large market is costly. A market for a specific musical CD satisfies all of these three criteria. As CDs are typically inexpensive we can expect that first, the resale values are small, and second, the variation in individual valuations is significantly larger than any possible correlation between valuations of bidders. For this reason, the independent private value framework is a good approximation of this environment. Moreover, the markets for labels released more than 1 year ago are generally small (especially if we do not look at the albums of extremely popular groups and singers), but the existing fan community provides a steady demand for such CDs. Last, as old releases are usually inexpensive and the market size is generally no more than 50 listed items, inflating such market and sustaining the larger market size for a period of 2-3 weeks becomes possible.

In this paper we use the dataset which is collected for the field experiment analysis in Nekipelov (2007). We look at bidding on eBay for the Robbie Williams' CD "The Greatest Hits", which was released in 2004. The singer is popular mostly in Europe, Australia and New Zealand. However, a fair fraction of Robbie Williams' CDs sells to the United States. In the market before the experiment, we observe roughly 24 simultaneously listed items (taking average across time). This relatively small number of listed items implies that it is possible to artificially double the market supply by listing additional items. In addition to this, the market before the intervention is quite active and sellers very rarely have to relist the CDs because of the absence of bids. The active feature of the market allowed us to collect more than 150 observations during one month before the market expansion. The observations before the market expansion form the control dataset. Then by listing additional items we managed to keep the market size of 60 items per day for a period of 3 weeks. The observations collected from the inflated market form the treatment dataset, where the treatment is market expansion.

Table 1: Characteristics of the auctions for "Greatest hits" CDs in the control group

variable	N obs.	mean	st. dev.	min	max
Picture (yes=1)	136	0.8823529	0.3233808	0	1
Duration	136	7.632353	1.919911	3	10
Shipping cost ($)	136	7.269701	6.490984	0	37.2
Seller's feedback score	136	13414.46	22421.4	0	164546
Store (yes=1)	136	0.6029412	0.4910972	0	1
Condition	136	0.6911765	0.4637162	0	1
# of bids	136	1.875	3.009707	0	10
# of early bids	136	0.1102941	0.6957308	0	6
# of bidders	136	1.279412	1.95388	0	8

The control dataset in our study is a set of auctions for the same CDs when the market was not inflated. In total we have collected the data for 136 auctions with the earliest auction starting on August 23, 2006 and the last auction ending on October 4, 2006. 15 auctions had items located in North America, 69 auctions in Europe (predominantly the UK), 43 auctions in Asia (mainly China and Taiwan), and 9 auctions in Australia. The shipping cost can be substantial if the item is located very far from the winning bidder. In these circumstances the shipping costs might play an important role in bidding behavior and we will use regional indicators to capture this effect in the data.

The original dataset contains auction-specific variables such as picture dummy, duration (in days), shipping cost, the percentage of positive feedback of the seller, seller's feedback score, and a dummy equal to 1 if the seller has a store. We also report auction characteristics such as the number of bids, number of bidders and number of early bids. We formally define an early bid as a bid submitted within the first 80% of auction duration and with an increment more than 10% in the final price in the auction. The descriptive statistics of the variables in the control group are presented in Table 1.

To build the treatment dataset we listed additional items and kept the market size of 60 items per day for 3 consecutive weeks (the market was inflated for several more consecutive weeks as we were relisting the items which were not sold, but the corresponding observations were discarded). We were listing CDs which we have previously purchased directly from a wholesale record store, all of which are original licensed copies. To make sales less dependent on the seller's feedback score, and not to raise bidder's suspicion about the authenticity of CDs we used 5 different sellers' accounts with different feedback score and listed from 5 to 10 CDs on each account. We chose auction duration time of 7 or 10 days as these are the most frequently used durations of auctions in the analyzed market. We sold 60 CDs in total. As in some cases we had to relist CDs, in total we created more than 100 auctions. We conducted the experiment from October 4, 2006 to November 10, 2006. This made the data collection period sufficiently far away from Thanksgiving and Christmas sales, and provided adequate homogeneity between the market conditions in the control group of observations and in the treatment group. To avoid the "transition effects" in the beginning, and to take into account we did not have enough CDs to sustain the market size at 60 listed items for more than 4 weeks, we only used the data from 20 days of the experiment. This allowed us to build the treatment dataset of 156 auctions with the earliest auction starting on October 8, 2006 and the last auction ending on October 28, 2006. The auctions in the treatment dataset are as diversely located as the auctions in the control dataset. However, as the number of items that we listed was substantial, the proportion of auctions located in the US is bigger in the treatment dataset than that in the control dataset. The dataset contains 70 items located in North America (these were dominated by the items listed from our seller's accounts), 56 auctions in Europe, 25 auctions in Asia, and 5 auctions in Australia. The basic statistics of the treatment dataset are given in Table 2.

Table 2: Characteristics of the auctions for "Greatest hits" CDs in the treatment group

variable	N obs.	mean	stdev	min	max
Picture (yes=1)	156	0.8205128	0.3849957	0	1
Duration	156	7.685897	2.23406	3	10
Shipping cost ($)	156	4.924754	2.894402	1.59	18.1
Seller's feedback score	156	34448.39	65604.64	0	176262
Store (yes=1)	156	0.4615385	0.5001241	0	1
Condition	156	0.9102564	0.2867346	0	1
# of bids	156	0.6858974	1.751698	0	12
# of early bids	156	0.25	0.9060727	0	7
# of bidders	156	0.5192308	1.127207	0	7

3.3 Estimation results

In the empirical analysis, we assume that the model satisfies the conditions which we impose on our theoretical econometric model. In particular, the population moment condition 2 is assumed to hold.

We look at a linear model of dependence between the standard deviation of bids in the auction STD and the early bidding dummy EARLY:

$$\mathrm{STD} = \beta_0 \mathrm{EARLY} + x'\beta + \epsilon.$$

We define any bid submitted in the first half of the auction as early. In our analysis we compare three different models. The first model is a simple OLS regression, in which we expect to obtain a coefficient for the early bidding dummy that is contaminated by the attenuation bias caused by omission of the unobserved visibility of the auction. The second model is 2SLS in which we use the set of observed regressors (other than early bidding dummy) x and the binary dummy indicating the treatment group of auctions as instruments Z. The third model is the efficient semiparametric method of moment model based on optimal non-linear vector of instruments that is developed in Theorem 1 for the subset of compliers.

To clarify the new GMM estimation procedure that we use, we will describe the steps to obtain the estimator in more detail. First, note that for the subset of compliers, the early bidding dummy and the vector of regressors are orthogonal to the residual:

$$e = \text{STD} - \beta_0 \text{EARLY} - x'\beta.$$

This means that for the subset of compliers a consistent estimate of β_0 can be obtained in a modified OLS framework. Abadie, Angrist, and Imbens (2002) suggest that a population moment condition can be modified to hold for the set of compliers if it is multiplied by:

$$k\left(\text{EARLY}, Z, x\right) = 1 - \frac{\left(1 - \text{EARLY}\right) Z}{\mathcal{Q}(x)} - \frac{\text{EARLY}\left(1 - Z\right)}{1 - \mathcal{Q}(x)}.$$

It is also noted in Abadie, Angrist, and Imbens (2002) that simple multiplication of the moment equation by $k\left(\cdot\right)$ may lead to a non-convex objective function. To create a convex model we substitute Z by $v_z\left(\text{STD}, \text{EARLY}, x\right) = E\left[Z \mid \text{STD}, \text{EARLY}, x\right]$ and create a new factor $\widetilde{k}(\cdot)$. We use this procedure to obtain first-stage estimates. We estimate $\mathcal{Q}(x)$ and $v_z\left(\text{STD}, \text{EARLY}, x\right)$ by running a logit regression on Hermite polynomials of the regressors up to the third degree. We then substitute the estimates to form the moment equation:

$$\widehat{m}^1\left(\beta\right) = \frac{1}{n}\sum_{i=1}^{n} \widetilde{\kappa}\left(\text{STD}, \text{EARLY}, x\right) e_i \begin{pmatrix} \text{EARLY}_i \\ x_i \end{pmatrix}.$$

The first-stage estimate of β is obtained by minimizing a standard GMM objective function.

At the second stage, we estimate the mean and the variance of the dependent variable $\text{STD} \mid \text{EARLY}, z, x$ by kernel smoothing. Specifically, given the first stage (inefficient) estimate of parameter β, we evaluate the elements of the weighting matrix. We estimate estimate

$$\hat{\omega}_{\text{EARLY},z}\left(x, \bar{\beta}\right) = \widehat{V}\left(\text{STD} | \text{EARLY}, z, x\right)$$

and $\hat{\gamma}_{\text{EARLY},z}\left(x, \bar{\beta}\right) = \widehat{E}\left(\text{STD}|\text{EARLY}, z, x\right) - \bar{\beta}_0 \text{EARLY} - \bar{\beta}' x$ for $\text{EARLY} = 0, 1$ and $z = 0, 1$. Non-parametric estimates are implemented through a local linear regression with normal kernel. In particular we evaluate $\widehat{E}\left(\text{STD}|\text{EARLY}, z, x\right)$ by minimizing objective:

$$\frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{\xi - \xi_i}{h}\right)\left[\text{STD} - \xi_i' \mu_\xi\right]^2,$$

where $\xi = (\text{EARLY}, z, x)$. Then

$$\widehat{E}\left(\text{STD}|\text{EARLY}, z, x\right) = \xi' \hat{\mu}_\xi.$$

Using weighting by the conditional probability of EARLY, Z given x we obtain the efficient estimate. Then we form the empirical moment $\widehat{m}\left(\beta\right)$ using non-parametric estimates of the relevant conditional distributions and form the efficient weighting matrix based on the first-stage estimates. In particular, we set up the exactly identified moment vector

$$\hat{m}\left(\beta_0, \beta\right) = \frac{1}{n}\sum_{i=1}^{n}\left(\widehat{\mathcal{P}_1(x_i)} - \widehat{\mathcal{P}_0(x_i)}\right)\begin{bmatrix} 1 & 0 \\ x_i & x_i \end{bmatrix}\widehat{\widehat{\Omega}}^{-1}(x_i)\begin{pmatrix} \frac{\text{EARLY}_i}{\widehat{\mathbf{P}}(x_i)} \\ \frac{1-\text{EARLY}_i}{1-\widehat{\mathbf{P}}(x_i)} \end{pmatrix}\left[\text{STD}_i - \beta_0\text{EARLY}_i - x_i'\beta\right].$$

We use a non-linear equation solver to recover parameters β_0 and β by finding the solution to $\hat{m}\left(\beta_0, \beta\right) = 0$. To evaluate the variance, we use the asymptotic formula for the semiparametric efficiency bound:

$$\widehat{\text{V}}\left(\hat{\beta}_0, \hat{\beta}\right) = \frac{1}{n}\sum_{i=1}^{n}\left(\widehat{\mathcal{P}_1(x_i)} - \widehat{\mathcal{P}_0(x_i)}\right)^2\begin{bmatrix} 1 & 0 \\ x_i & x_i \end{bmatrix}\widehat{\widehat{\Omega}}^{-1}(x_i)\begin{bmatrix} 1 & x_i' \\ 0 & x_i' \end{bmatrix}.$$

The estimation results for OLS, 2SLS, and GMM models are presented in Table 3. We use the market expansion dummy and the starting price as instruments for the early jump bidding dummy. One can see that the coefficient for the early bidding dummy is significant and positive in both OLS and 2SLS. This means that early jump bids tend to increase the variance of later bids in the auction. However, the coefficient estimated with 2SLS is almost twice the size of that in the simple linear model. This reflect a significant downward bias in the OLS estimate, which is consistent with our prediction of omitted variable bias.

Using the model from Abadie, Angrist, and Imbens (2002) we reestimate the model for a subsample of auctions for which the market expansion has significantly shifted the entry pattern (and thus, influenced early bidding behavior). Compliers in our population are the auctions for which the market expansion has certainly shifted the entry patterns (and thus, they would now have early bids if the market had not been inflated). The results of this estimator is shown in the last column of Table 3. In the efficient GMM procedure, we used the staring values from the inefficient one. With our choice of the bandwidth in

Table 3: Dependence between variance of bids and early jump bid dummy

	OLS	2SLS	LATE(ineff.)	LATE (eff.)
Early bid	.4398	1.353	1.2459	1.2459
	(.1301)***	(.3663)***	(.3340)***	(.1241)***
Condition	-.1659	-.0258	-.7426	-.7426
	(.1072)	(.1372)	(0.4840)	(.1576)***
Seller's	-.0136	-.0111	-.0313	-.0313
feedback score	(.0040)***	(.0056)**	(.0433)	(.0241)
Europe dummy	.2059	.1464	.5815	.5815
	(.0727)***	(.0796)*	(.5409)	(.2815)**
Australia dummy	.3052	.1770	-.1542	-.1542
	(.1566)*	(.1089)	(.2723)	(.1823)
Constant	.1883	.0005	.7096	.7096
	(.1051)*	(.1438)	(.2587)**	(.2022)**
Observations	292	292	292	292
$\mathbf{R}^2$	.2896	.4678	-	-

Robust standard errors in parentheses * significant at 5% level; ** significant at 1% level non-parametric evaluation of moments, both procedures lead to the same point estimates[3]. However, one can see that the efficient GMM produces tighter standard errors. The effect of early bidding on the variance of bids in the auction for the subsample of compliers is comparable with the effect the entire sample with the binary treatment instrument. This is consistent with the endogeneity of the binary dummy.

4 Conclusion

In this paper we provide an efficient estimation framework for a linear moment-based semi-parametric model of treatment effects and apply this model to the analysis of the dependence between the variance of bids on eBay and early bidding. Unlike sealed-bid or button auctions, a feature of the multi-auction environment of eBay is that early bidding tends to increase the variance of bids. A regression analysis of this relationship will not provide unbiased estimate of the effect of the binary treatment because both the variance of bids and early bidding are effected by the unobserved heterogeneity in entry of bidders into auctions. Our estimation procedure is based on the availability of a binary instrument for the the endogenous regressor in the field experiment data on eBay where the supply in the market for s certain pop music CD was doubled. Theoretical analysis predicts that such influence should change the early bidding incentive for bidders. However, it should not change the structure of the unobserved heterogeneity across auctions. As a result, we can use the indicator for the increased market size as a binary instrument. We find the semiparametric efficiency bound for this model and provide a GMM-based procedure which allows us to obtain an estimator which achieves the semiparametric efficiency bound. Our empirical results find that ignoring the effect of unobserved heterogeneity in the auction data leads to a substantial bias in the estimate of the effect of early bidding on the variance of bidding.

[3] We observe that bandwith variation influences both standard errors and the estimates in the efficient GMM procedure. We report the estimates where the bandwith was chosen as $\alpha \mathrm{Var}(X)^{0.5}$ and α was picked through cross-validation.

References

ABADIE, A., J. ANGRIST, AND G. IMBENS (2002): "Instrumental Variables Estimates of the Effects of Subsidized Training on the Quantiles of Trainee Earnings," *Econometrica*, 70, 91–117.

ANGRIST, J. (2004): "Treatment Effect Heterogeneity in Theory and Practice," *Economic Journal*, 114(494), 83.

CHERNOZHUKOV, V., AND C. HANSEN (2005): "An IV Model of Quantile Treatment Effects," *Econometrica*, 73(1), 245–261.

DUNFORD, N., AND J. T. SCHWARZ (1958): *Linear Operators. Part I: General Theory.* Wiley.

FIRPO, S. (2003): "Efficient Semiparametric Estimation of Quantile Treatment Effects," *Working Paper, UC Berkeley Department of Economics.*

FROLICH, M. (2006): "Nonparametric IV estimation of local average treatment effects with covariates," *Journal of Econometrics*, 139(1), 35–75.

HAHN, J. (1998): "On the Role of the Propensity Score in Efficient Semiparametric Estimation of Average Treatment Effects," *Econometrica*, 66(2), 315–331.

HIRANO, K., G. IMBENS, AND G. RIDDER (2003): "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score," *Econometrica*, 71(4), 1161–1189.

HONG, H., AND D. NEKIPELOV (2007): "Semiparametric Efficiency in Nonlinear LATE Models," working paper, Stanford University and Duke University.

NEKIPELOV, D. (2007): "Entry Deterrence and Learning Prevention on eBay," Job market paper, Duke University.

NEWEY, W. (1990): "Efficient Instrumental Variables Estimation of Nonlinear Models," *Econometrica*, 58(4), 809–837.

Proof of Theorem 1

Consider the parametrization θ for the model with covariates x. Let $\phi(x)$ be the density of x with the support on $\mathcal{X}$ (or the probability mass function for the discrete components of x), the joint likelihood function for the data along a parametric path of the model can be factorized as:

$$f_\theta\left(w,z,x\right)=\left[f_\theta^1\left(w_1 \mid w_2,\, z,\, x\right)\right]^{w_2}\,\left[f_\theta^0\left(w_1 \mid w_2,\, z,\, x\right)\right]^{(1-w_2)}\mathcal{F}_\theta^{w_2}\left(1-\mathcal{F}_\theta\right)^{(1-w_2)}$$
$$\times\mathcal{Q}_\theta^z(x)\left(1-\mathcal{Q}_\theta(x)\right)^{(1-z)}\,\phi(x).$$

where $f_\theta^k\left(w_1|w_2,z,x\right)=f_\theta\left(w_1|w_2=k,z,x\right)$. The score of the model can be written as:

$$S_\theta\left(w,z \mid x\right)=(1-w_2)s_\theta^0\left(w_1 \mid w_2,\, z,\, x\right)+w_2 s_\theta^1\left(w_1 \mid w_2,\, z,\, x\right)$$

$$+\frac{(1-z)\dot{\mathcal{P}}_{0\theta}(x)}{\mathcal{F}(z,x)\left(1-\mathcal{F}(z,x)\right)}\left[w_2-\mathcal{F}(z,x)\right]+\frac{z\dot{\mathcal{P}}_{1\theta}(x)}{\mathcal{F}(z,x)\left(1-\mathcal{F}(z,x)\right)}\left[w_2-\mathcal{F}(z,x)\right]$$

$$+\frac{\dot{\mathcal{Q}}_\theta(x)}{\mathcal{Q}(x)\left(1-\mathcal{Q}(x)\right)}\left[z-\mathcal{Q}(x)\right]+s_\theta(x),$$

where $s_\theta(x)$ is the score corresponding to the distribution of x. The expression for the tangent set of the model for conditional distribution moments will be similar to that of a model with unconditional moment equations:

$$\mathcal{T}=\left\{(1-w_2)s_\theta^0\left(w_1 \mid w_2,\, z,\, x\right)+w_2 s_\theta^1\left(w_1 \mid w_2,\, z,\, x\right)+z\,\xi(x,z)\left[w_2-\mathcal{F}(z,x)\right]\right.$$

$$\left.+(1-z)\,\zeta(x,\, z)\left[w_2-\mathcal{F}(z,x)\right]+a(x)\left[z-\mathcal{Q}(x)\right]+t(x)\right\},$$

where $E_\theta\left[s_\theta^i\left(w_1 \mid w_2,\, z,\, x\right) \mid w_2,\, z,\, x\right]=0$ for $i=0,1$, $E\{t(x)\}=0$ and $\zeta(\cdot)$, $\xi(\cdot)$ and $a(\cdot)$ are square - integrable functions.

Now consider the directional derivative of the parameter vector β determined by the conditional moment equation $\varphi\left(x,\, w_2,\, \beta\right)$ in equation (4). We assume that the support of x - the set $\mathcal{X}$ is non-degenerate. In this case we can potentially identify a parameter vector β with arbitrarily many dimensions. We assume here that $\beta=(\beta_0,\beta_1)$ is k - dimensional. Our strategy now will be to define a linear functional A which will transform the conditional moment equation to an exactly identified system of unconditional moments on $\mathbb{R}^k$. Such a functional can be defined by instrument functions $\mathcal{A}(x,\, w_2) \,:\, \mathcal{X} \times \{0,\, 1\} \mapsto \mathbb{R}^k$ such that for function $f(x,\, w_2)$:

$$A \circ f = E\left[\mathcal{A}(x,\, w_2)\, f(x,\, w_2)\right]. \tag{5}$$

We assume that $\varphi(\beta, x, w_2)$ and $\frac{\partial \varphi(\beta,x,w_2)}{\partial \beta}$ are elements of $\mathbf{L}_2(\mathcal{X} \times \{0,1\},\, \mathfrak{F},\, \mu)$ for each $\beta \in \mathcal{B}$, where $\mathfrak{F}$ is a Borel σ -algebra in the product space $\mathcal{X} \times \{0,1\}$ and μ is the corresponding probability measure. In this case the dual space to $\mathbf{L}_2(\mathcal{X} \times \{0,1\},\, \mathfrak{F},\, \mu)$ will also be $\mathbf{L}_2(\mathcal{X} \times \{0,1\},\, \mathfrak{F},\, \mu)$ (see Dunford and Schwarz (1958)). As a result, limiting our analysis to the functions $\mathcal{A}(\cdot)$ in $\mathbf{L}_2(\mathcal{X} \times \{0,1\},\, \mathfrak{F},\, \mu)$ will allow us to find the optimal structure of the unconditional moment vector. In fact, as it is shown in Dunford and Schwarz (1958) any continuous linear functional on $\mathbf{L}_p(T,\, \mathfrak{F},\, \mu)$ can be represented as (5) for $\mathcal{A} \in \mathbf{L}_q(T,\, \mathfrak{F},\, \mu)$ for $p^{-1} + q^{-1} = 1$ and $1 < p < \infty$.

We assume that the probability measure on $\mathcal{X} \times \{0,1\}$ does not depend on the parameter β. Then define the Jacobi matrix:

$$J = E\left[\mathcal{A}(x, w_2)\, \frac{\partial \varphi(\beta, x, w_2)}{\partial \beta}\right] = -E\left[\mathcal{A}(x, w_2)\left(w_2,\, x'\right)\right].$$

Consider a parametrization path θ of our semi-parametric model. Along the parametrization path we can derive the directional derivative of the parameter β as:

$$J\, \frac{\partial \beta}{\partial \theta} = A \circ \left[g(w, x, \beta)\, s_*(w_1 \mid w_2, x)\right],$$

where $s_*(\cdot)$ is the score of the model corresponding to the chosen parametrization path.

The score can be expressed as

$$s_*(w_1 \mid w_2, x)\, f_*(w_1 \mid w_2, x)$$

$$= \frac{w_2\mathcal{P}_1(x) + (1-w_2)(1-\mathcal{P}_0(x))}{\mathcal{P}_1(x)-\mathcal{P}_0(x)} s_\theta(w_1 \mid w_2, z=w_2, x)\, f(w_1 \mid w_2, z=w_2, x)$$

$$- \frac{w_2\mathcal{P}_0(x) + (1-w_2)(1-\mathcal{P}_1(x))}{\mathcal{P}_1(x)-\mathcal{P}_0(x)} s_\theta(w_1 \mid w_2, z=1-w_2, x)\, f(w_1 \mid w_2, z=1-w_2, x)$$

$$+ \left[\frac{[1-w_2+\mathcal{P}_1(x)(2w_2-1)]\,\dot{\mathcal{P}}_{0\theta} - [1-w_2+\mathcal{P}_0(x)(2w_2-1)]\,\dot{\mathcal{P}}_{1\theta}}{(\mathcal{P}_1(x)-\mathcal{P}_0(x))^2}\right]$$

$$\times \left(f(w_1 \mid w_2, z=w_2, x) - f(w_1 \mid w_2, z=1-w_2, x)\right).$$

The Jacobi matrix is non-singular for almost all $\mathcal{A}(\cdot)$. Then introducing:

$$\widetilde{g}(w,x,\beta) = J^{-1}\mathcal{A}(x,w_2)\left(w_1 - \beta_0 w_2 - x'\beta_1\right),$$

we can express the directional derivative of β as:

$$\frac{\partial\beta(\theta)}{\partial\theta} = E\left\{\int \widetilde{g}(w,x,\beta)\, s_*(w_1 \mid w_2, x)\, f_*(w_1 \mid w_2, x)\, dw_1\right\}.$$

The efficient influence function $\Psi(w,z,x)$ is an element of the tangent set such that:

$$\frac{\partial\beta(\theta)}{\partial\theta} = E\left[\Psi(w,z,x) S_\theta(w,z,x)\right].$$

Denote

$$\begin{aligned}\widetilde{\Delta}(w_2,x) &= \{E[\widetilde{g}(w,x,\beta) \mid w_2, z=1, x] - E[\widetilde{g}(w,x,\beta) \mid w_2, z=0, x]\} \\ &= J^{-1}\mathcal{A}(x,w_2)\left(E[w_1 \mid w_2, z=1, x] - E[w_1 \mid w_2, z=0, x]\right).\end{aligned}$$

Then we can find the efficient influence function corresponding to the structure of the tangent set for our model as:

$$
\begin{aligned}
\Psi(w,z,x) \quad = \quad & \frac{\mathbf{P}(x) w_2\, z}{\mathcal{Q}(x)\left(\mathcal{P}_1(x)-\mathcal{P}_0(x)\right)}\left(\widetilde{g}(w,x,\beta)-E\left[\widetilde{g}(w,x,\beta)\mid w_2=1,z=1,x\right]\right) \\
- \quad & \frac{\mathbf{P}(x) w_2\,(1-z)}{\left(1-\mathcal{Q}(x)\right)\left(\mathcal{P}_1(x)-\mathcal{P}_0(x)\right)}\left(\widetilde{g}(w,x,\beta)-E\left[\widetilde{g}(w,x,\beta)\mid w_2=1,z=0,x\right]\right) \\
+ \quad & \frac{\left(1-\mathbf{P}(x)\right)\left(1-w_2\right)\,(1-z)}{\left(1-\mathcal{Q}(x)\right)\left(\mathcal{P}_1(x)-\mathcal{P}_0(x)\right)}\left(\widetilde{g}(w,x,\beta)-E\left[\widetilde{g}(w,x,\beta)\mid w_2=0,z=0,x\right]\right) \\
- \quad & \frac{\left(1-\mathbf{P}(x)\right)\left(1-w_2\right)\, z}{\mathcal{Q}(x)\left(\mathcal{P}_1(x)-\mathcal{P}_0(x)\right)}\left(\widetilde{g}(w,x,\beta)-E\left[\widetilde{g}(w,x,\beta)\mid w_2=0,z=1,x\right]\right) \\
+ \quad & \frac{\widetilde{\Delta}(w_2=1,x)\mathbf{P}(x)}{\left(\mathcal{P}_1(x)-\mathcal{P}_0(x)\right)^2}\left[\frac{\mathcal{P}_1(x)(1-z)}{1-\mathcal{Q}(x)}-\frac{\mathcal{P}_0(x)z}{\mathcal{Q}(x)}\right]\left(w_2-\mathcal{F}(z,x)\right) \\
+ \quad & \frac{\widetilde{\Delta}(w_2=0,x)\left(1-\mathbf{P}(x)\right)}{\left(\mathcal{P}_1(x)-\mathcal{P}_0(x)\right)^2}\left[\frac{\left(1-\mathcal{P}_1(x)\right)(1-z)}{1-\mathcal{Q}(x)}-\frac{\left(1-\mathcal{P}_0(x)\right)z}{\mathcal{Q}(x)}\right]\left(w_2-\mathcal{F}(z,x)\right).
\end{aligned}
$$

The semiparametric efficiency bound is equal to the variance of the efficient influence function. Note that the vector $\mathcal{A}\left(w_2,x\right)$ can be represented as:

$$
\mathcal{A}\left(w_2,x\right)=\left(\mathcal{P}_1(x)-\mathcal{P}_0(x)\right)\left(\frac{\mathcal{Q}(x)w_2}{\mathbf{P}(x)}+\frac{\left(1-\mathcal{Q}(x)\right)\left(1-w_2\right)}{1-\mathbf{P}(x)}\right)\mathcal{M}(x)\left(\begin{array}{c}\frac{w_2}{\mathbf{P}(x)}\\ \frac{1-w_2}{1-\mathbf{P}(x)}\end{array}\right),
$$

where $\mathcal{M}(x)$ is a $k\times 2$ matrix (k is the size of the Euclidean parameter β). Denote

$$
D(x)=\operatorname{diag}\left\{\frac{\mathcal{Q}(x)}{\mathbf{P}(x)},\ \frac{1-\mathcal{Q}(x)}{1-\mathbf{P}(x)}\right\},\quad \text{and}\quad \theta\left(x\right)=\left(\begin{array}{c}\frac{w_2}{\mathbf{P}(x)}\\ \frac{1-w_2}{\mathbf{P}(x)}\end{array}\right)\left(w_2,x'\right)'.
$$

Using these notations, the Jacobi matrix can be written as

$$
J=E\left\{\left(\mathcal{P}_1(x)-\mathcal{P}_0(x)\right)\mathcal{M}(x)\left(\begin{array}{c}\frac{\mathcal{Q}(x)w_2}{\mathbf{P}^2(x)}\\ \frac{(1-\mathcal{Q}(x))(1-w_2)}{(1-\mathbf{P}(x))^2}\end{array}\right)\left(w_2,\ x'\right)\right\}=E\left\{\left(\mathcal{P}_1(x)-\mathcal{P}_0(x)\right)\mathcal{M}(x)D(x)\theta(x)\right\}.
$$

Using the notation for $\overline{\Omega}$ that we have introduced in the Theorem, we can express the variance of the efficient influence function as:

$$\begin{aligned}
V\left(\widehat{\beta}\right) &= J^{-1}E\left\{\mathcal{M}(x)D(x)\overline{\Omega}(x)D(x)\mathcal{M}(x)'\right\}J^{-1\prime} \\
&= E\left\{\left(\mathcal{P}_1(x)-\mathcal{P}_0(x)\right)\mathcal{M}(x)D(x)\theta(x)\right\}^{-1} \\
&\qquad \times E\left\{\mathcal{M}(x)D(x)\overline{\Omega}(x)D(x)\mathcal{M}(x)'\right\}E\left\{\left(\mathcal{P}_1(x)-\mathcal{P}_0(x)\right)\mathcal{M}(x)D(x)\theta(x)\right\}^{-1\prime}.
\end{aligned}$$

It can be verified that the minimum variance is achieved when

$$\mathcal{M}(x) = \left(\mathcal{P}_1(x)-\mathcal{P}_0(x)\right)\theta(x)'\overline{\Omega}(x)^{-1}D(x)^{-1},$$

and the semiparametric efficiency bound is

$$V\left(\widehat{\beta}\right) = E\left\{\left(\mathcal{P}_1(x)-\mathcal{P}_0(x)\right)^2\theta(x)'\overline{\Omega}(x)^{-1}\theta(x)\right\}^{-1}.$$

The result of Theorem 1 follows from substituting . the expression for $\theta(x)$ into this formula.