EMPIRICAL IMPLEMENTATION OF NONPARAMETRIC FIRST-PRICE AUCTION
MODELS

Daniel J. Henderson
John A. List
Daniel L. Millimet
Christopher F. Parmeter
Michael K. Price

Empirical Implementation of Nonparametric First-Price Auction Models
Daniel J. Henderson, John A. List, Daniel L. Millimet, Christopher F. Parmeter, and Michael
K. Price
NBER Working Paper No. 17095
May 2011
JEL No. C12,C14,D44

ABSTRACT

Nonparametric estimators provide a flexible means of uncovering salient features of auction data."Although
these estimators are popular in the literature, many key features necessary for proper implementation
have yet to be uncovered. Here we provide several suggestions for nonparamteric estimation of first-price
auction models. Specifically, we show how to impose monotonicity of the equilibrium bidding strategy;
a key property of structural auction models not guaranteed in standard nonparametric estimation. We
further develop methods for automatic bandwidth selection. Finally, we discuss how to impose monotonicity
in auctions with differering number of bidders, reserve prices, and auction-specific characteristics.
Finite sample performance is examined using simulated data as well as experimental auction data.

Daniel J. Henderson
Dept. of Economics
State University of New York at Binghamton
Binghamton, NY 13902-6000
djhender@binghamton.edu

John A. List
Department of Economics
University of Chicago
1126 East 59th
Chicago, IL 60637
and NBER
jlist@uchicago.edu

Daniel L. Millimet
  Southern Methodist University
Department of Economics
Box 0496
Dallas, TX 75275-0496
and IZA
millimet@mail.smu.edu

Christopher F. Parmeter
University of Miami
cparmeter@bus.miami.edu

Michael K. Price
Department of Economics
University of Tennessee
515 Stokely Management Center
Knoxville, TN 27996
and NBER
mprice21@utk.edu

1 Introduction

Nonparametric kernel methods are advantageous in many empirical settings because they are largely immune, relative to parametric alternatives, to concerns regarding mis-specification of functional forms. In light of this, as well as the logic set forth in McAfee and Vincent (1992) suggesting that policy conclusions should hinge on distribution-free methods when knowledge of the underlying distribution is vacuous, nonparametric methods have thus become a cornerstone of structural estimation of auctions. Despite this popularity, however, nonparametric methods possess theoretical and computational deficiencies; notably, the 'curse of dimensionality' and derivations of asymptotic properties. Yet, even ignoring these unavoidable issues, other empirical issues remain unresolved in the nonparametric estimation of first-price auction models.

In this paper, we address two of these issues: bandwidth selection and imposition of smoothness constraints (e.g., monotonicity). We do so in the context of the seminal first-price auction estimator of Guerre, Perrigne, and Vuong (2000; GPV hereafter); although the methods proposed should also extend naturally to other auction settings. Moreover, the increasing popularity of identifying and constructing nonparametric, structural microeconometric estimators should benefit from the discussion here.

In terms of bandwidth selection, we are unaware of any automated selection procedures that have been proposed in the structural auction literature. Even though bandwidth selection mechanisms are well-studied in nonparametric density and regression estimation, and are considered of the utmost importance, there is a noticeable void in the first-price auction setting (Athey and Haile 2008). There is a good explanations for this void: the values of interest are unobserved, thus making it difficult to formulate a criterion for smoothing the equilibrium bidding strategy. Nonetheless, we attempt to fill this void by proposing a least-squares cross-validation procedure to obtain bandwidths. In contrast to the widely used rule-of-thumb estimates employed in the first-price auction literature, which attempts to provide an optimal bandwidth for the bid density, we provide a framework for automated selection of the bandwidth for the bid-value relationship. The intuition for the proposed approach is identical to that in the density estimation context. While the density values are unobserved (as are the underlying values in our setting), the use of a leave-one-out estimator helps to circumvent this lack of important information when constructing the squared error criterion. We note that the optimal bandwidth for the bid density may not be suitable for the bid-value relationship. In fact, in our empirical example, we show how using a rule-of-thumb bandwidth can lead to poor estimates.

The second issue we address concerns the imposition of smoothness constraints such as monotonicity. In constrast to parametric alternatives, the imposition of such constraints to ensure theoretically-consistent estimates in nonparametric settings is not straightforward. In an empirical structural auction setting, this is a primary concern as smoothness conditions are implied by the underlying structure of the equilibrium bidding behavior. Moreover, the underlying theory for the GPV first-price auction estimator is but one of a number of theoretical reasons why the monotonicity assumption is important

in a structural model, among them being the 'single-crossing' property recognized in Athey (2001) and Athey and Haile (2002, 2008).

Our approach involves tilting the empirical distribution of the data by the least amount, relative to a benchmark, to achieve the desired smoothness constraints (monotonicity in our case). This approach, known as constraint weighted bootstrapping (Hall, Huang, Gifford, and Gijbels, 2001; Hall and Huang, 2001), is firmly entrenched in the statistics literature. Furthermore, the constrained approach proposed here will produce a monotonically-constrained estimator regardless of the bandwidth deployed. Finally, our proposed constrained estimator has the nice feature that it is numerically identical to the GPV estimator when the latter is monotonic. Nonetheless, it is generally the case that the constrained estimator will differ from that of GPV in small sample settings.

Beyond the general case of imposing smoothness constraints in a nonparametric setting, we discuss how to impose monotonicity in auctions with differing numbers of bidders, reserve prices, and auction-specific characteristics. We further discuss bandwidth selection with these features and believe that our paper provides necessary tools to perform proper empirical applications of structural models. We showcase our methods in simple settings using both simulated and experimental data. We find that there are often finite sample gains when resorting to constrained estimation and that the constrained and unconstrained estimators tend towards one another as the sample size increases. Specifically, our simulated evidence shows that imposing monotonicity can produce improved estimates of the bid-value relationship as well as policy relevant variables, such as (but not limited to) the optimal reserve price. Moreover, using experimental auction data we show how the rule-of-thumb bandwidth can easily be misleading relative to our data-driven approach.

Prior to continuing, it is worth re-emphasizing that our estimator should also be of interest not only in an auction setting, but also in other areas where structural approaches are employed to recover the primitives of economic models. For example, the ability to constrain nonparametric estimators should prove indispensable for those who wish to use monotone comparative statics (e.g., Athey 2001, 2002) to recover structural parameters in stochastic optimization problems or games of incomplete information. We also show that the semiparametric, optimal reserve price estimator of Li, Perrigne, and Vuong (2003) can be improved in certain dimensions when monotonicity is imposed on the bid-value relationship.

More generally, constrained nonparametric estimation has important implications within economics as a whole, as there are many instances where economic theory provides some information about the model being estimated (e.g., Hotz and Miller 1993; Olley and Pakes 1996). In addition to these applications, the approach described here should be of interest to econometricians interested in constrained nonparametric methods (Gallant 1981; Matzkin 1994) and our approach complements recent work that examines the imposition of curvature conditions in nonparametric settings (Beresteanu 2007; Chak, Madras, and Smith 2005; Chernozhukov, Fernandez-Val, and Galichon 2009; Du, Parmeter, and Racine 2010). The constraint weighted bootstrapping approach advocated here is similar in spirit to the empirical likelihood methods developed in Owen (1988) and the information theoretic approaches to GMM

presented in Imbens, Spady, and Johnson (1998). In addition to the papers discussed above, a substantial amount of consideration has been paid to the issue of monotonicity by statisticians; see the citations in Henderson and Parmeter (2009).

The remainder of the paper is laid out as follows. In Section 2 we briefly review the theory underlying the first-price auction setup within the independent private value paradigm (IPVP) as well as the nonparametric estimator proposed by GPV. Section 3 discusses the importance of the bandwidth on the GPV estimator and proposes a data-driven method to select the bandwidth. In Section 4 we describe how to implement the constraint weighted bootstrapping theory developed in the statistics literature to create a generalized estimator that imposes monotonicity. Additionally, we discuss how to extend this methodology to a variety of auction settings likely to arise in practice. Section 5 provides a small simulation study to illustrate that the method performs well when monotonicity is violated, as well as simulations to showcase how estimation of the optimal reserve price can be improved when monotonicity is imposed. A formal application of the method to experimental first price auction data that was recently used to adjudicate opposing structural estimators is also presented to highlight the importance of which bandwidth is used in practice. In Section 6 we emphasize the usefulness of this style of nonparametric estimation beyond structural auctions and indicate several lines of possible future research.

2 Theoretical Background and Estimation

2.1 Preliminaries

Within the IPVP each player knows his or her value of the product to be auctioned, v, but no other player's value. Players' values are assumed to be independent draws from $F(v)$ (the CDF), which is taken as common knowledge. Players select their bidding strategy to maximize their expected payout, given by $\pi^e(\cdot)$. This leads to the following maximization problem:

$$\max_b \quad \pi^e(b) = (v-b)F(\sigma_n)^{n-1}, \tag{1}$$

where b is the player's corresponding bid when there are n total participants in the auction and $\sigma_n = \beta_n^{-1}(b)$ denotes the inverse of the bid function, $\beta_n(v)$, used by the player. The first order condition is given by:

$$-F(\sigma_n)^{n-1} + (n-1)(v-\beta_n)F(\sigma_n)^{n-2}f(\sigma_n)\sigma_n' = 0. \tag{2}$$

The assumption that the bid function is monotonic implies $d\sigma_n/db = 1/\beta_n'(v)$. Furthermore, with symmetry of the bidders, $\beta(v) = b$. These features allow us to simplify the solution to

$$\beta_n' + \frac{(n-1)f(v)}{F(v)}\beta_n = \frac{(n-1)vf(v)}{F(v)}, \tag{3}$$

4

which is a linear differential equation with solution given by

$$\beta_n(v) = v - \frac{\int\limits_{\underline{v}}^{v} F(u)^{n-1} du}{F(v)^{n-1}}, \tag{4}$$

assuming the absence of a reserve price,[1], where $\underline{v}$ represents the minimum of the support of the value distribution. If we allow for a reserve price, then the differential equation has the solution

$$\beta_n(v) = v - \frac{\int\limits_{r}^{v} F(u)^{n-1} du}{F(v)^{n-1}} \quad \text{where} \quad r \leq v. \tag{5}$$

Assuming that all potential bidders place bids, the only difference between equations (4) and (5) are the limits of integration. In essence the reserve price acts as a boundary condition in exactly the same way that $\underline{v}$ does in the no reserve setting. Paarsch and Hong (2006) provide a more detailed description of this derivation and the IPVP in general.

2.2 Nonparametric Estimation in First-Price Auctions

In the seminal paper on the nonparametric structural identification and estimation of a first-price auction, GPV provide a natural setting in which to think about the distribution of valuations within the IPVP in a nonparametric framework. Their analysis spurred (perhaps started) the growth of nonparametric structural estimation of auctions across paradigms, including affiliated private values (Li, Perrigne, and Vuong 2002), unobserved heterogeneity with independent values (Krasnokutskaya 2009) and conditionally independent private information (Li, Perrigne, and Vuong 2000). We describe their method under the situation of no reserve price.

The structural equilibrium bidding strategy derived in GPV is given as

$$v_i = b_i + \frac{G(b_i)}{(n-1)g(b_i)} = \xi(b_i, n, G), \tag{6}$$

where v_i and b_i are the value and bid for agent i, respectively. $G(b_i)$ is the cumulative distribution function (CDF) of the bid density and $g(b_i)$ is the probability density function (pdf) of bids. Only the vector of bids is observed by the econometrician. Once the functional forms of $G(\cdot)$ and $g(\cdot)$ are assumed or estimated, then the values (v_i) can be estimated along with the corresponding CDF and pdf, $F(\cdot)$ and $f(\cdot)$, respectively.

The nonparametric estimation approach given in GPV is as follows:

---

[1]A reserve price is such that all submitted bids must be greater than this price.

1. Estimate $g(b)$ using kernel methods.

$$\widehat{g}(b) = \frac{1}{nTh} \sum_{i=1}^{n} \sum_{t=1}^{T} K\left(\frac{b - b_{it}}{h}\right), \tag{7}$$

where $b_i t$ is the bid for agent i in auction t. Thus, we are pooling bids from multiple auctions with the identical number of bidders to increase the sample size. $K(\cdot)$ is a kernel function required to be a non-negative, bounded and symmetric density function with compact support. The bandwidth, h, depends on the sample size and converges to zero as T goes to infinity. The standard bias-variance tradeoff exists when considering the choice of the bandwidth.

2. Estimate $G(b)$ using the empirical CDF

$$\widehat{G}(b) = \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \mathbf{1}\{b_{it} \leq b\}, \tag{8}$$

where $\mathbf{1}\{A\}$ is the indicator that the event A is true.

3. Construct $\widehat{v}_{it} = \xi(b_{it}, n, \widehat{G})$ using the above estimates to recover the values.

4. Estimate the density and distribution of values, $f(\widehat{v}_{it})$ and $F(\widehat{v}_{it})$, using equations (7) and (8) above with the same bandwidth, with b_{it} and b replaced with $\widehat{v}_{it}$ and v, respectively.

The above discussion omits three important details. First, the GPV estimator needs to trim the sample near the boundaries of the pseudo-values, $\widehat{v}_{it}$; see equation (6) on page 531 in GPV). Kernel density estimators are well known to be inconsistent near the edge of the support of the variable of interest. This contaminates the second-stage recovery of the distribution of values. GPV propose trimming observations that are within one bandwidth of $\underline{b}$ and $\bar{b}$, the upper and lower bounds of the support for bids. This yields a consistent estimator on the interior of $\mathcal{I} = [\underline{b}, \bar{b}]$. Second, no formal method is provided to select the bandwidth used for smoothing the bid-value function even though it is well known in applied nonparametric research that the bandwidth employed can have a serious impact on the results. Section 3 provides a data-driven bandwidth selection technique that is similar in spirit to the common data-driven methods used for bandwidth selection in the kernel density estimation literature. Third, GPV show that $\xi(b_i, n, G)$ is strictly increasing for all $b_{it} \in \mathcal{I}$ (condition $C2$ of Theorem 1 in GPV). Their nonparametric approach, however, does not formally impose this condition in the estimation. To fill this void, Section 4 proposes a method to impose this monotonicity condition.

3 Bandwidth Choice

3.1 Monotonicity and Bandwidth Selection

To highlight the importance of bandwidth choice, we begin by showing that there exists a bandwidth such that any estimated equilibrium bidding strategy is monotonic. Formally, our claim is that if $\mathcal{I}$

is a compact interval, then for all sufficiently large bandwidth h, $\widehat{\xi}'(\cdot|h) > 0$ on $int(\mathcal{I})$, guaranteeing monotonicity of the estimated equilibrium bidding strategy. To justify this statement, we assume that the kernel function, $K(\cdot)$, has two continuous derivatives in a small neighborhood of the origin, with $K'(0) = 0$.[2] We note that as $h \to \infty$, the following two relations hold:

$$K\left(\frac{b - b_{it}}{h}\right) \approx K(0) + \left(\frac{b - b_{it}}{h}\right) K'(0) + o_p(h^{-2})$$

and

$$K'\left(\frac{b - b_{it}}{h}\right) \approx h^{-1}K'(0) + \left(\frac{b - b_{it}}{h^2}\right) K''(0) + o_p(h^{-3}).$$

Together, these relations imply

$$\widehat{g}(b|h) \approx h^{-1}K(0)$$

and

$$\widehat{g}'(b|h) \approx h^{-3}K''(0) \cdot (nT)^{-1} \sum_{i=1}^{n} \sum_{t=1}^{T} (b - b_{it}) + o_p(h^{-3})$$

uniformly over $b \in \mathcal{I}$ since $h \to \infty$ implies that $|b_{it} - b|/h \to 0$. From this it follows that $\widehat{g}(b)^2 > |\widehat{g}'(b)|$ $\forall\, b \in \mathcal{I}$ for sufficiently large h. However, consistent nonparametric estimation requires that h shrinks to zero as n gets large. Hence, we cannot choose a very large h to achieve a monotone estimation result asymptotically.

This shows that the numerator of the derivative of the GPV estimator will be positive everywhere given a sufficiently large bandwidth; the denominator is always positive. This is intuitive since $h \to \infty$ implies that $\hat{g}(b)$ is flat for all b and thus the derivative is zero everywhere.[3]

This result is discouraging since bandwidth selection has, up to this point, been arbitrary in this setting. While asymptotically optimal bandwidth are available for a given sample size and kernel, they are at best 'asymptotically' optimal. Alternatively, the smallest bandwidth that guarantees monotonicity could be employed. However, such arbitrary selection of the bandwidth is not suggested. Further, choosing an arbitrarily large bandwidth to guarantee monotonicity may mask interesting and essential information about a density. In general, most economic applications require automated criteria for bandwidth choice and a separate means by which to impose monotonicity.

3.2 Automated Bandwidth Selection

Automated bandwidth selection in structural auction settings has been largely unexplored even though these issues have taken place under the mantle of nonparametric density and regression estimation (Athey and Haile 2008). It might be posited that automated selection has not been addressed because

---

[2]This property holds for all standard kernels used in the auction literature (Gaussian, Epanechnikov, and triweight).

[3]An alternative proof may be given by noting that $\hat{\xi}'(\cdot) = 1 + (n-1)^{-1} + (n-1)^{-1}\left(\hat{G}(\cdot)\hat{g}'(\cdot)/\hat{g}(\cdot)^2\right)$, which implies that $\hat{g}(\cdot)$ becomes flat as $h \to \infty$, meaning $\hat{g}'(\cdot) \to 0$ so that $\hat{\xi}'(\cdot)$ is positive everywhere for large h. We are indebted to an anonymous referee for this point.

the values, v, are unobserved, making it difficult to formulate a criterion for smoothing of the equilibrium bidding strategy. Instead, researchers often resort to rule-of-thumb bandwidths or asymptotic bandwidth selection criteria which are based on assumptions of the underlying distribution.

GPV use the triweight kernel, $K(u) = (35/32)(1 - u^2)^3 \mathbf{1}(|u| \leq 1)$, and recommend the rule-of-thumb bandwidth $1.06\widehat{\sigma}_b(NT)^{-1/5}$, which is optimal for estimating the density of bids if they were normally distributed. Of first note is that this bandwidth actually leads to undersmoothing, as $(4/3)^{1/5}$ is the asymptotically optimal scale factor for the Gaussian kernel, not the triweight kernel. The asymptotically optimal scale factor for the triweight kernel is $2.978 \cdot (4/3)^{1/5} = 3.154$. This scaling factor comes from the theory on canonical kernels found in Marron and Nolan (1989). Essentially, to guarantee that the same degree of smoothing is present when different kernels are used, the bandwidth must be adjusted by a specific factor, in our case 2.978.[4] Li, Perrigne, and Vuong (2003) advocate for this rescaled bandwidth and apply it when studying optimal reserve price estimation.

However, even with an appropriately modified rule-of-thumb bandwidth, it is desirable to have data driven methods to obtain the appropriate smoothness. For example, if the density of bids were log-normally normally distributed (as in Paarsch, 1992) with unit variance as opposed to normally distributed, the optimal scale factor would then be 0.7069, which is nearly 4.5 times smaller than the bandwidth for normally distributed data. Given the result in the previous sub-section, this is one possible explanation for why monotonicty is often found in empirical studies. A separate point of note is that it is not clear that selection of the bandwidth based on optimal estimation for the density of bids will necessarily be suitable for estimation of the bid-value relationship.

Although there exist many bandwidth selection criteria in both the density and regression settings, least-squares cross-validation (LSCV) is a popular empirical choice. The idea behind the method is to minimize the integrated square error of the estimator. For a given estimator $\widehat{f}$ of an unknown density f, the integrated square error (ISE) can be written as $\int \left(\widehat{f}(x) - f(x)\right)^2 dx$. An analog estimator is easily constructed noting that the $\int f(x)^2 dx$ is independent of the bandwidth. While this approach automates the selection of the bandwidth in order to optimally smooth the bid density, it is unknown whether this is the appropriate amount of smoothing for the equilibrium bidding strategy. Further, if we were to smooth the CDF instead of using the empirical distribution function to estimate $\widehat{G}(\cdot)$, it is not clear how to select a bandwidth when both the distribution and density are smoothed simultaneously (i.e., using the same bandwidth).

Here, we argue that given the primary aim of most structural work is to recover the underlying pseudo-values to perform policy experiments that will enable the researcher to identify the optimal reserve price and/or auction format, the objects of interest to dictate appropriate smoothing are the pseudo-values, $\widehat{v}_{it} = \widehat{\xi}(b_{it}, n, G)$, used to construct the density of values, a bandwidth which is optimal for the density of bids may not be adequate for construction of the pseudo-values. As these are our objects of interest, we choose to minimize the LSCV function focusing on the estimation of $\xi(\cdot)$. It may

---

[4]See equation 2.5 in Marron and Nolan (1989).

seem that an appropriate estimator of the bandwidth for $\widehat{\xi}(b_{it}, n, G)$ would be to compare the squared differences between $\xi(b, n, G)$ and $\widehat{\xi}(b, n, G)$ over all feasible bids. However, given that we do not observe $\xi(b, n, G)$ it would seem reasonable to replace it with a leave-one-observation or leave-one-bidder out estimate. Unfortunately, in this context the bandwidth that minimizes this function is $h = \infty$ because the differences between the actual estimator and the leave-one-out estimator diminish as the smoothness increases.

In light of this, we propose to using the bandwidth which minimize the sample analog of the ISE,

$$
\begin{aligned}
ISE(h) &= \int \left[\widehat{\xi}(b, n, G) - \xi(b, n, G) \right]^2 db \\
&= \int \widehat{\xi}(b, n, G)^2 \, db - 2 \int \widehat{\xi}(b, n, G) \, \xi(b, n, G) \, db + \int \xi(b, n, G)^2 \, db.
\end{aligned}
\tag{9}
$$

The last term is independent of the bandwidth and so we focus on the construction of analog versions of the first two terms in our decomposition of ISE. For the unknown $\xi(b, n, G)$ appearing in the second term, we replace it with a leave-one-bidder-out estimator instead of a leave-one-observation-out estimator (i.e., we are using $\widehat{\xi}_{-i}(b_{it}, n, G)$ instead of $\widehat{\xi}_{-it}(b_{it}, n, G)$).[5] This gives us

$$
\begin{aligned}
ISE(h) &= \int \widehat{\xi}(b, n, G)^2 \, db - 2 \int \widehat{\xi}(b, n, G) \, \xi(b, n, G) \, db \\
&= \int \left[b^2 + \frac{2b\hat{G}(b)}{(n-1)\hat{g}(b)} + \frac{\hat{G}(b)^2}{(n-1)^2 \hat{g}(b)^2} \right] db \\
&\quad - 2 \int \left[b^2 + \frac{b\hat{G}(b)}{(n-1)\hat{g}(b)} + \frac{b\hat{G}_{-i}(b)}{(n-1)\hat{g}_{-i}(b)} + \frac{\hat{G}_{-i}(b)^2}{(n-1)^2 \hat{g}_{-i}(b)^2} \right] db \\
&= \int \left[-b^2 + \frac{\hat{G}(b)^2}{(n-1)^2 \hat{g}(b)^2} - \frac{2b\hat{G}_{-i}(b)}{(n-1)\hat{g}_{-i}(b)} - \frac{2\hat{G}_{-i}(b)^2}{(n-1)^2 \hat{g}_{-i}(b)^2} \right] db.
\end{aligned}
\tag{10}
$$

The first term in $ISE(h)$ also does not depend upon the bandwidth and so our final decomposition of the integrated squared error becomes

$$
ISE(h) = \int \left[\widetilde{\xi}(b, n, G)^2 - 2b\widetilde{\xi}_{-i}(b, n, G) - \widetilde{\xi}_{-i}(b, n, G)^2 \right] db,
\tag{11}
$$

where $\widetilde{\xi}(b, n, G) = \frac{\hat{G}(b)}{(n-1)\hat{g}(b)}$ and $\widetilde{\xi}_{-i}(b, n, G) = \frac{\hat{G}_{-i}(b)}{(n-1)\hat{g}_{-i}(b)}$. Data-driven bandwidths can then be obtained via

$$
\min_h \sum_{i=1}^{n} \sum_{t=1}^{T} \left[\widetilde{\xi}(b_{it}, n, G)^2 - 2b_{it}\widetilde{\xi}_{-i}(b_{it}, n, G) - \widetilde{\xi}_{-i}(b_{it}, n, G)^2 \right].
\tag{12}
$$

The benefit of using estimated values is that these bandwidths also control the smoothing of the density of the bids. However, it is not naïve to the fact that we must also calculate the distribution of the bids and then use both of these estimates to obtain the values. Hence, we are able to use a data driven

---

[5]This is a common approach when dealing with panel type data (e.g., Henderson, Li, and Carroll 2008; Kneip and Simar 1996).

approach to obtain estimates of our objects of interest.

This type of approach could be used to select bandwidths in other auction settings where the values are not observed. This would include the nonparametric estimator of the affiliated private value auction proposed in Li, Perrigne, and Vuong (2002). In fact, for any of the nonparametric estimators proposed that involve the standard first order differential equation as in GPV equation (3), it is hypothesized that a similar type of LSCV criteria could be constructed to obtain data-driven bandwidths focused on estimation of the unknown values.

4 Monotone Estimation of the Bid Function

4.1 Baseline Case

Once an automated procedure is used to choose the bandwidth, there is no guarantee that the GPV estimator will produce a monotonic result. Rather than reverting to parametric methods in such instances, we instead show how a modified version of the GPV estimator can be constrained to be monotonically increasing. To do so, we utilize the constraint weighted bootstrap technique from the statistics literature.

In this literature, it is becoming commonplace to impose monotonicity when estimating survival functions. Noting that $\xi(\cdot)$ is similar to a survival function, our approach is as follows:

1. Estimate $g(b)$ as

$$\widehat{g}(b|p) = \frac{1}{h} \sum_{i=1}^{n} \sum_{t=1}^{T} p_{it} K\left(\frac{b - b_{it}}{h}\right),$$ (13)

   where the p_{it} are observation-specific weights. Note, the GPV estimator is a special case of our estimator where each weight is set equal to $1/nT$.[6]

2. Estimate $G(b)$ as

$$\widehat{G}(b|p) = \int_{-\infty}^{b} \widehat{g}(u|p)du.$$ (14)

Thus, we are not constructing the CDF of the bids using the empirical distribution function. To ensure that our CDF corresponds to the pdf estimated in equation (13), we need to integrate the pdf as opposed to simply estimating the CDF by the empirical distribution function. This step is not done in GPV, nor is it a common approach in studies that use both a CDF and a pdf in the estimation.[7] One may think that the reason for this is twofold. First, the asymptotic arguments are most likely easier to prove given widely known properties of the empirical distribution estimator.

---

[6]The terminology constraint weighted bootstrap stems from the fact that we are introducing probability weights to shift the true observations so that the constraint(s) of interest are satisfied. This is consistent with the notion of biased bootstrapping. Even though no resampling is occurring, the weights act as though the data have been resampled in a manner which delivers the constrained surface. We mention here that this methodology also has intimate links with empirical likelihood. We elect to use the terminology constraint weighted bootstrapping for the remainder of our discussion.

[7]See Martins-Filho and Yao (2008) for a recent example that does.

Second, the empirical distribution estimator is easier to construct than an integral of an estimated probability density.[8]

3. Construct $\widehat{v}_{it} = \widehat{\xi}(b_{it}, n, G|p)$ using equation (6) to recover the values. Employ the truncation strategy of GPV (page 531, equation (6)).

4. Estimate the density and distribution of values, $f(\widehat{v}_{it})$ and $F(\widehat{v}_{it})$, using equations (7) and (8) above, with $\widehat{v}_{it}$ and v in the place of b_{it} and b, respectively. Note, because we have a two-step estimator, the recovery of the density and distribution of the values does not require the incorporation of any constraint weights. This is because monotonicity has been imposed on the estimator of the bid function which is used to create the pseudo-values; the resulting pseudo-values can then be treated as they are in GPV.

The crucial feature of our estimator is that the weights, p_{it}, are selected to ensure that the estimated values are monotonically increasing in the bids. To select the vector of weights, we choose $p = \{p_{11}, p_{12}, \ldots, p_{1T}, p_{21}, \ldots, p_{nT}\}$ to minimize a distance metric subject to the constraint that $\widehat{\xi}'(b_{it}, n, G|p) \geq 0$ on $\mathcal{I}$. If we desire to impose strict monotonicity, $\widehat{\xi}'(b_{it}, n, G|p) > 0$ on $\mathcal{I}$, then we need to pick some small number δ such that $\widehat{\xi}'(b_{it}, n, G|p) > \delta$ on $\mathcal{I}$ so that this becomes computationally feasible.[9] We also impose the regularity conditions $p_{it} \geq 0 \ \forall i, t$ and $\sum_{i=1}^{n} \sum_{t=1}^{T} p_{it} = 1$. These conditions make the weights act as though they are drawn from a density and will prove useful when making comparisons to the uniform weights, $1/nT$, used in GPV. For simplicity, we choose to impose our nonnegativity constraint on

$$\hat{\xi}_1 = n\widehat{g}(b|p)^2 - \widehat{G}(b|p)\widehat{g}'(b|p). \tag{15}$$

Noting that $\hat{\xi}'/\hat{\xi}_1$ is always nonnegative, this implies that both have the same sign.

Our distance metric is the power divergence measure introduced in Cressie and Read (1984) and proposed in Hall, Huang, Gifford, and Gijbels (2001) for monotone estimation of a hazard rate.[10] The power divergence measure is

$$D_\rho(p) = \frac{1}{\rho(1-\rho)}[nT - \sum_{i=1}^{n} \sum_{t=1}^{T} (nTp_{it})^\rho], \quad -\infty < \rho < \infty. \tag{16}$$

where $\rho \neq 0, 1$. We need to take limits for $\rho = 0$ or 1. They are given as

$$D_0(p) = -\sum_{i=1}^{n} \sum_{t=1}^{T} \log(nTp_{it}); \quad D_1(p) = \sum_{i=1}^{n} \sum_{t=1}^{T} p_{it} \log(nTp_{it}). \tag{17}$$

---

[8]This does not preclude the use of kernel estimation of the CDF. See Bowman, Hall, and Prvan (1998) or Li and Racine (2007) for a discussion of CDF estimation via kernel methods. Additionally, this approach may allow us to avoid numerical integration all together as the integration sign can be brought inside the summations and typically $\int_{-\infty}^{b} K(u)du$ is known. We thank an anonymous referee for making this computational connection.

[9]This is also suggested in Hall and Huang (2001).

[10]It is also used in Hall and Huang (2001) for nonparametric monotone estimation of a regression function.

A value of $\rho = 0.5$ corresponds to Hellinger distance. Note, for all ρ we have $D_\rho(p) \geq 0 \;\forall p$ and $D_\rho(p) = 0$ if and only if $p_{it} = 1/nT \;\forall i, t$. This suggests that departures from uniformity of the weights will correspond to a positive divergence measure, indicating the presence of regions of non-monotonicity.

Regardless of the sampling distribution for the values of the players across auctions and the choice of $\mathcal{I}$, it is entirely plausible that $\widehat{\xi}_1$ will have a zero crossing on $\mathcal{I}$. An example can easily be constructed where a point lies both a bandwidth away from the boundary and from its nearest point. The probability of this event is strictly positive given minimal assumptions about the bid density. Fortunately, data sets that produce this event or a similar event are pathological in nature. Even if we cannot find a set of weights that guarantees a monotonic estimator, this should pose no problem. In fact, we can view this event as providing information about the true equilibrium bidding strategy or as evidence that other features of the auction are being ignored by the econometrician (Athey and Haile 2008).

While it may be argued that this procedure is entirely heuristic given the fact that many papers have confirmed monotonicity between bids and values, the ability to easily impose this condition when estimating models using auction data is important from an economic standpoint.[11,12] Indeed, even if the weights are uniform, the researcher can be confident that the estimated equilibrium bidding strategy is monotonic. This is more formal than visual inspection of the estimated surface. Additionally, while it may appear that monotonicity holds unconditionally between bids and values, the presence of covariates renders visual inspection useless in higher dimensions.

Theoretically, this estimator (ignoring truncation) is consistent following only minor modifications of the proof in Hall, Huang, Gifford, and Gijbels (2001). Given that the pseudo-values are constructed identically to GPV, the theoretical properties of the value density estimator should follow directly. We do not consider asymptotic normality of this estimator and leave that for future research.

4.2 Heterogeneous Auctions and Reserve Prices

Many auctions are characterized by differing numbers of bidders, the use of reserve prices, and auction-specific heterogeneity. We discuss an extension to the baseline case encapsulating all of these features to highlight the ease by which the constraint weighted bootstrap technique discussed above may be generalized. The GPV estimator in this auction setting relies on the following first order condition (written in terms of the actual bids):

$$v_{it} = b_{it} + \frac{1}{\mathcal{N} - 1} \left\{ \frac{G(b_{it}|X_t)}{g(b_{it}|X_t)} + \frac{H(r_t|X_t)}{[1 - H(r_t|X_t)]g(b_{it}|X_t)} \right\}, \tag{18}$$

where $\mathcal{N}$ is the number of potential bidders, r_t is the reserve price in auction t, and X_t is a $d \times 1$ vector of auction-specific observables.

---

[11]We note that this confirmed monotonicity may have been arrived at by employing a rule-of-thumb bandwidth that was too large.

[12]See Figure 2 of Li, Perrigne, and Vuong (2000) for an example where $\widehat{\xi}(\cdot)$ is locally but not globally monotonic in outer continental shelf wildcat auctions.

Within the IPVP, the number of actual bidders in auction t, n_t, has a binomial distribution with parameters $\mathcal{N}$ and $1 - H(r_t|X_t)$. A natural candidate estimator for $\mathcal{N}$ is (GPV; Paarsch and Hong 2006)

$$\widehat{\mathcal{N}} = \max_{t=1,\ldots,T} n_t. \tag{19}$$

Our approach to estimating $H(r_t|X_t)$ follows GPV and uses the fact that $H(r_t|X_t) = 1 - E[n_t|X_t]$. To that end, we use the standard local-constant kernel estimator

$$\hat{H}(r_t|X_t) = 1 - \frac{1}{\widehat{\mathcal{N}}T(h_1 \cdots h_d)} \sum_{t=1}^{T} n_t A_t(x), \tag{20}$$

where

$$A_t(x) = \frac{K_h(x, X_t)}{\sum_{t=1}^{T} K_h(x, X_t)} \tag{21}$$

and h_l, $l = 1, \ldots, d$, is the bandwidth associated with the lth element in X.

Before proceeding to estimation, we mention that the observed bids, b_{it}, must be transformed as $s_{it} = \sqrt{b_{it} - r_t}$ (see GPV) due to the proportionality between the actual bid density, $g(b)$, and $1/\sqrt{b - r}$ as b approaches r. This transformation prevents the density of bids from becoming unbounded near the reserve price. Using this transformation, we can write the first order condition in equation (18) as

$$v_{it} = r_t + s_{it}^2 + \frac{2s_{it}}{\mathcal{N} - 1} \left\{ \frac{G^*(s_{it}|X_t)}{g^*(s_{it}|X_t)} + \frac{H(r_{it}|X_t)}{(1 - H(r_{it}|X_t))g^*(s_{it}|X_t)} \right\} = \xi(s_{it}, \mathcal{N}, G^*, r_t, X_t), \tag{22}$$

where $G^*(\cdot)$ and $g^*(\cdot)$ are the CDF and pdf, respectively, of the transformed bids. The GPV estimator then follows by pooling bids across auctions and estimating the unobserved valuations using the following algorithm:

1. Estimate $\mathcal{N}$ and $H(r_t|X_t)$ as indicated above. The bandwidths used to construct $H(r_t|X_t)$ can be obtained via standard least-squares cross-validation.

2. Estimate $g^*(s|x)$ as
$$\widehat{g}^*(s|p, x) = \frac{1}{h_s} \sum_{t=1}^{T} \sum_{i=1}^{n_t} p_{it} K\left(\frac{s - s_{it}}{h_s}\right) A_t(x). \tag{23}$$

3. Estimate $G^*(s|x)$ as
$$\widehat{G}^*(s|p, x) = \sum_{t=1}^{T} \sum_{i=1}^{n_t} p_{it} \tilde{K}\left(\frac{s - s_{it}}{h_s}\right) A_t(x), \tag{24}$$

where $\tilde{K}$ is the corresponding CDF kernel. For example, the CDF kernel corresponding to the triweight kernel, $K(u) = (35/32)(1 - u^2)^3 \mathbf{1}(|u| \leq 1)$, is

$$\tilde{K}(u) = (35/32)(1 + u)^4(16/35 - (29/35)u + (4/7)u^2 - (1/7)u^3)\mathbf{1}(|u| \leq 1). \tag{25}$$

Notice that we are not integrating the estimated pdf as we did in the baseline auction setting. Due to the smoothing of the auction covariates, we can select a CDF kernel whose derivative is equivalent to the kernel we use to construct our conditional pdf. Thus, we have exactly the definition of a conditional pdf, $f(u|w) = \frac{\partial F(u|w)}{\partial u}$.

4. Construct $\widehat{v}_{it} = \widehat{\xi}(s_{it}, \mathcal{N}, G^*, r_t, X_t|p)$ using the above estimates to recover the values following the truncation strategy of GPV (page 550).

5. Estimate the density and distribution of values, $f(\widehat{v}_{it}|X_t)$ and $F(\widehat{v}_{it}|X_t)$, as in GPV (page 550).

To minimize the computational burden in selecting the weights, we choose to impose our non-negativity constraint on

$$\widehat{\xi}_1 = 2s\widehat{\mathcal{N}}\widehat{g}^*(s|p,x)^2 + s\widehat{G}^*(s|p,x)\left(1 - 2\widehat{g}^{*\prime}(s|p,x)\right) + \widehat{H}(r|x)s(\widehat{g}^*(s|p,x) - 2)/(1 - \widehat{H}(r|x)). \tag{26}$$

We also maintain the regularity conditions $p_{it} \geq 0 \ \forall i, t$ and $\sum_{i=1}^{n}\sum_{t=1}^{T} p_{it} = 1$, consistent with use of the power divergence metric.

Finally, this approach could be extended to allow for differing reserve prices across auctions or to map the arguments here to other auction settings such as the affiliated private values paradigm where Li, Perrigne, and Vuong (2002) developed a similar nonparametric estimator. Again, their estimator must be monotonic in the bids (see their Proposition 1). Applying our methodology to this estimator is straightforward given that it has the same form, but involves a multivariate density as opposed to a univariate one.

Following our earlier approach in the simple auction setting, a bandwidth can be obtained for the bids in the heterogenous auction setting. This would follow from:

$$LSCV(h_s) = \int \left(\hat{\xi}(s, \mathcal{N}, G^*, r, X) - \xi(s, \mathcal{N}, G^*, r, X)\right)^2 ds.$$

This estimator represents a trivial extension beyond the homogeneous auction setting given that the bandwidths for the covariates are obtained prior to construction of $\hat{\xi}(\cdot)$. However, an alternative strategy would be to select the bandwidths for X simultaneously with the smoothing parameter for the bids. This can be handled by minimizing $LSCV(h)$ over both h_s and the vector of smoothing parameters for the covariates.

4.3 Optimal Reserve Price

The optimal reserve price is one of the few pieces of an auction that a seller can manipulate (outside of improving the quality of the product being auctioned) to increase revenues. This optimal reserve price is important to learn from a policy standpoint. Let $f(v)$ continue to denote the pdf of values for the bidders with CDF $F(v)$. The seller places value v_0 on the good to be auctioned. The optimal reserve price, r^*, is determined by maximizing total expected revenue (utility), which for an auction with n

bidders is

$$\Pi(r) = v_0 F(r)^n + n \int_r^{\bar{v}} \left[v f(v) + F(v) - 1 \right] F(v)^{n-1} dv \tag{27}$$

where $\bar{v}$ represents the maximum of the support of $f(v)$.[13] The first term in $\Pi(r)$ is the expected value the seller obtains if no bids are placed which exceed r. The second term is the expected revenue from each bidder submitting a bid higher than r so that the object is sold.

To find the optimal reserve price, we differentiate total expected revenue with respect to r to obtain[14]

$$n v_0 F(r)^{n-1} f(r) - n \left[r f(r) + F(r) - 1 \right] F(r)^{n-1}(r) = 0. \tag{28}$$

Solving (28) for the optimal reserve yields:

$$r^* = v_0 + \frac{1 - F(r^*)}{f(r^*)}. \tag{29}$$

While this is a simple and elegant solution to determining the optimal reserve price, it depends upon the unknown distribution and density of values. In empirical settings, only the bids are observed.

Recently, Li, Perrigne and Vuong (2003) propose a semiparametric estimation method to determine the optimal reserve price that depends upon the observed bids as opposed to the unobserved values. We briefly describe their method here. As is typical, we will have observations for T auctions and will observe bids b_{it} for bidder i in auction t. Let $B_{it} = \max_{j \neq i} b_{jt}$ denote the largest bid in auction t not equal to the bid by agent i. Corollary A.1 in Li, Perrigne, and Vuong (2003) shows that the optimal reserve price satisfies $r^* = \xi(x_0)$ where x_0 maximizes expected profit, defined as:

$$\Pi(x) = E \left[v_0 1 \left\{ B_1 \leq x \right\} 1 \left\{ b_1 \leq x \right\} \right.$$
$$\left. + n \left[b_1 + (\xi(x) - x) \left(\frac{G(x)}{G(b_1)} \right)^{n-1} \right] 1 \left\{ B_1 \leq b_1 \right\} 1 \left\{ b_1 \geq x \right\} \right]. \tag{30}$$

Note that for any given reserve price x, the seller's expected profit is composed of the likelihood that the reserve price is higher than any bids submitted, resulting in the value of the good being auctioned off as the profit, while the second piece is the value in excess of the initial value of the good that results from the reserve price being set so that qualifying bids are placed generating additional value. x_0 is the value which will maximize the overall sum of these pieces.

---

[13]We assume here that the reserve price does not cause bidders to exit the auction.

[14]The derivative of the integral was determined via Leibniz's rule:

$$\frac{d}{d\alpha} \int_{a(\alpha)}^{b(\alpha)} f(x, \alpha) dx = \frac{db(\alpha)}{d\alpha} f(b(\alpha), \alpha) - \frac{da(\alpha)}{d\alpha} f(a(\alpha), \alpha) + \int_{a(\alpha)}^{b(\alpha)} \frac{\partial}{\partial \alpha} f(x, \alpha) dx.$$

Replacing $\xi(x)$ with its formal definition in (6) we have

$$
\begin{aligned}
\Pi(x) =& E\left[v_0 1\left\{B_1 \leq x\right\} 1\left\{b_1 \leq x\right\}\right. \\
&\left.+n\left[b_1 + \left(\frac{G(x)}{(n-1)g(x)}\right)\left(\frac{G(x)}{G(b_1)}\right)^{n-1}\right] 1\left\{B_1 \leq b_1\right\} 1\left\{b_1 \geq x\right\}\right] \\
=& E\left[v_0 1\left\{B_1 \leq x\right\} 1\left\{b_1 \leq x\right\}\right. \\
&\left.+n\left[b_1 + \left(\frac{G(x)^n}{(n-1)g(x)G(b_1)^{n-1}}\right)\right] 1\left\{B_1 \leq b_1\right\} 1\left\{b_1 \geq x\right\}\right].
\end{aligned}
\tag{31}
$$

The insight of Li, Perrigne, and Vuong (2003) is to replace $G(x)$ and $g(x)$ with nonparametric estimates. Thus, determination of x_0 requires an analog estimator for expected profit. Using the density and distributional estimators proposed earlier, we can average over all bids in all auctions to construct an estimated profit function for which the maximum can be easily determined given the one dimensional nature of the problem. We have

$$
\begin{aligned}
\widehat{\Pi}(x) = \frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T & \left(v_0 1\left\{B_{it} \leq x\right\} 1\left\{b_{it} \leq x\right\} + n b_{it} 1\left\{B_{it} \leq b_{it}\right\} 1\left\{b_{it} \geq x\right\}\right. \\
&\left.+\frac{n\hat{G}(x)^n}{(n-1)\hat{g}(x)\hat{G}(b_{il})^{n-1}} 1\left\{B_{it} \leq b_{it}\right\} 1\left\{b_{\max} - h \geq b_{it} \geq x\right\}\right).
\end{aligned}
\tag{32}
$$

Trimming is introduced for the last term since, as noted by Li, Perrigne and, Vuong (2003), $\hat{g}(x)$ is not well estimated near $\bar{b}$. Here, h is the bandwidth used to construct the kernel density estimator of $g(\cdot)$. Li, Perrigne, and Vuong (2003, Corollary A.2) show that any $\hat{x}_0 \in [b_{\min} + h, b_{\max} - h]$ which maximizes expected profit $\widehat{\Pi}(x)$, is a consistent estimator for x_0.

The beauty of determining the optimal reserve price is that this can be done after monotonicity of the equilibrium bid-value relationship is imposed. In this case, the weights can be used in the construction of both $\hat{g}(\cdot)$ and $\hat{G}(\cdot)$ prior to creating the expected profit function from which $\hat{x}_0$ will be determined.

5 Empirical Demonstration

To illustrate our approach, we begin by examining simulated data where we specify the exact form of the value distribution and create corresponding equilibrium bids. Next, we show that for certain distributions the monotonically-constrained GPV estimator can result in improved performance of the optimal reserve price in small settings. We then assess experimental data obtained from a laboratory setting to see if monotonicity arises in an artificial setting where bids are submitted by participants. This is useful because we know both the value distribution and the actual values, and can therefore more adequately address the criticisms of detecting monotonicity raised in Athey and Haile (2002, 2008).

5.1 Simulated Data

5.1.1 Estimation of the Bid-Value Relationship

Our simulation experiments examine monotonic distributions that are theoretically consistent with an equilibrium bidding strategy. We consider $T = 100$ auctions, each with $n = 5$ bidders, yielding 500 observed bids. We replicate each scenario 100 times.

We choose the true distribution of private values to be either log-normal with parameters zero and one or gamma with parameters one and three. For the log-normal case, we follow the truncation strategy in GPV, discarding those value draws that are below 0.055 and above 2.5. Similarly, for the gamma distribution, we discard values that are below 0.0455 and above 4.982. For every replication, we first draw nT values from the truncated distribution. We then compute the bids, b_{it}, using

$$b_{it} = v_{it} - \frac{1}{F(v_{it})^{n-1}} \int_{\underline{v}}^{v_{it}} F(u)^{n-1} du, \tag{33}$$

where $\underline{v}$ is the smallest value drawn from the truncated distribution for the given replication.

Using these generated data, we apply our estimation procedure to each replication. We use (13) and (14) to estimate the density and distribution of the bids for a given set of weights. We employ the triweight kernel with the rule-of-thumb bandwidth using the triweight scaling (as opposed to the 1.06 normal reference scaling) to speed up computing time. The weights are determined using $\rho = 0, 0.5$, and 1 and are found using the sequential quadratic programming routine SQPSolve in the programming language GAUSS 8.0. While our problem is not a quadratic programming problem, this type of solver uses a modified quadratic program to find the step length for moving in the direction of a minimum. As expected, each iteration takes longer to run than GPV. In addition to computation time issues, one problem that we encountered several times was that the program would not return feasible results. This was easily remedied, however, by changing the starting values.[15]

The simulation results are given in Figures 1 and 2. Panel (a) of each figure plots the true equilibrium bidding strategy along with the estimates from the GPV estimator and our estimator. The curves correspond to the 95th percentile of the distance metric. Panel (b) of each figure depicts the envelope-curves of the weights after the constraints have been achieved. It is clear that the true data generating process provides a monotonic equilibrium bidding strategy. However, the finite sample results of the GPV estimator show regions where the derivative is negative. Our estimator corrects for these regions of non-monotonicity by changing the weights. In Panel (b) of each figure we see that the corresponding weights deviate from $1/nT$ in the bid region where the GPV estimator is non-monotonic.

Of particular interest is whether or not the constrained estimator provides finite sample gains relative to the unconstrained GPV estimator. It turns out that the fit improves when the unconstrained function

---

[15]Our starting values were selected at random from a uniform distribution and divided by the sum of the starting values to preserve the summation constraint.

deviates further from monotonicity. For example, in the figures corresponding to the 95th percentile of the distance metric, the median value for the ratio of the absolute bias between the unconstrained and constrained estimators is greater than unity. Specifically, in the log-normal case, this ratio is 1.0011 when $\rho = 0$, 1.0013 when $\rho = 0.5$, and 1.0176 when $\rho = 1$. While this is a relatively minor improvement, in the gamma case the ratios are 1.1311, 1.2907 and 1.2145 for $\rho = 0$, 0.5, and 1, respectively. This limited evidence shows that the constrained estimator is at least as good as the unconstrained estimator and sometimes better in terms of bias.

5.1.2   Estimation of the Optimal Reserve Price

While it may be interesting from an econometric standpoint that the bias of the bid-value relationship is smaller with the constrained estimator, it may not be overly important from an economic standpoint. It is perhaps more important to discuss whether or not gains are made with respect to the estimation of a parameter of economic interest. Here, we consider estimation of the optimal reserve price by both the unconstrained and constrained estimators. Specifically, we again generate data from a log-normal and gamma distribution, but also consider data generated from a negatively skewed distribution, Beta(5,1). The purpose for the last distribution is that we expect the reserve price to lie in a region with fewer data points and hence we should get additional finite sample gains for the constrained estimator.

Table 1 presents the results of this exercise. We calculate the optimal reserve price as in Section 4.3 for both the unconstrained and constrained estimators. Given that Li, Perrigne, and Vuong (2003) prove theoretically that the estimates of the optimal reserve price are consistent as the number of auctions tends towards infinity, we start with relatively small numbers of auctions ($t = 5$ and 10). Even though the number of observations is relatively small in each run as compared to the previous sub-section, auctions of this size are common in the literature. We carry over the remaining settings from Section 5.1.1, but only report the results for $\rho = 0.5$ as the results for other values of this parameter did not qualitatively alter the results.

Although it is not obvious from the table, we did find that as the number of auctions increased, the bias, variance, and MSE decreased substantially. Hence, we find evidence to suggest that their asymptotic results hold. With respect to our table, we look at the median value for the ratio of the absolute bias, variance, and MSE between the unconstrained and constrained estimators. Values greater than unity suggest that the constrained estimator has finite sample gains.

Overall, the results are consistent with the fact that the constrained estimator can produce gains in the optimal reserve price estimator in small samples. However, we do not witness uniform gains across the three distributions because, as dictated by Theorem 4.3 of Hall and Huang (2001) and Theorem 2.2 of Du, Parmeter, and Racine (2010), outside of a diminishing interval, the GPV estimator and the constrained GPV estimator are exact. If the reserve price happens to fall in these areas, then the estimators produce identical results. What is interesting is that while there are minimal gains to estimating the bid-value relationship for the log-normal case in our earlier simulations, we see that the

bias of the optimal reserve price is considerably lower.

Examining the individual results, we find that each of the three different distributions gives a different result. For the log-normal distribution, there are large gains in bias at a cost of an increase in the variance for the constrained estimator. The MSE, however, is close to unity and each number tends towards unity as the number of auctions grows (as hypothesized). For the gamma distribution, there are improvements in variance at a (very minor) cost of bias and the MSE for the constrained estimator appears to be worse as compared to the unconstrained estimator, but this (and the other measures) tend towards one as the number of auctions increases. Finally, for the negative skewed, Beta (5,1)) distribution, we see improvements in bias, variance, and MSE, but these improvements are relatively small and the ratio tends towards unity as the number of auctions increases.

Table 1: Ratio of median bias, variance and MSE between unconstrained versus constrained estimator. A value above unity implies that the constrained estimator has a better finite sample performance at the median value.

	Log-Normal	Gamma(1,3)	Beta(5,1)
$n = 5, t = 10$			
Bias	1.2210	0.9888	1.0139
Variance	0.9513	1.0332	1.0092
MSE	0.9917	0.9782	1.0283
$n = 5, t = 20$			
Bias	1.1154	0.9897	1.0025
Variance	0.9734	1.0292	1.0016
MSE	1.0061	0.9793	1.0045

5.2   Experimental Data

Our experimental data were originally collected by Dyer, Kagel, and Levin (1989). Since the data are used in Bajari and Hortaçsu (2005) and are discussed there as well, we provide only limited details. MBA students at the University of Houston participated in a series of first-price sealed bid auctions over the course of two hours. Subjects submitted contingent bids based on the number of other bidders in the auction (either 2 or 5). However, we are treating the submitted bids (with either 3 or 6 bidders) as the actual bids for our purposes. Values were drawn from a $\mathcal{U}[0, 30]$ density. As in Bajari and Hortaçsu (2005), we drop the submitted bids for the first five auctions of a given run of the experiment. This leaves us with 23 auctions over three experimental runs. We have a total of 414 bids, regardless of the number of bidders, since we are ignoring the contingent bidding aspect of the experiment. In our analysis, we focus on the six bidder case.

We employ the triweight kernel as advocated in Guerre, Perrigne, and Vuong (2000) using both the rule-of-thumb bandwidth and the data-driven selection method described above. The rule-of-thumb

bandwidth is 7.044, whereas our method provides a bandwidth of 1.522. The gross difference in the bandwidths is most likely due to the fact that a rule-of-thumb bandwidth is arrived at assuming the unknown density is that of a normal random variate. However, given the experimental nature of the data, we know that the underlying value and bid distributions are both uniform. Thus, it is quite natural that our bandwidths differ.

The magnitude of the rule-of-thumb bandwidth produces a globally monotonic bid-value relationship (the dotted line in Panel (a)), whereas our data-driven method yields a non-monotonic relationship (the solid line in Panel (a)). This appears to be a case where previous research would suggest that a constrained estimator is unnecessary. However, we see that the ability to impose monotonicity here is important. Using $\rho = 0.5$, we obtain a distance metric of 0.870. Our estimated bid-value relationships are provided in Figure 3, Panel (a). The true relationship, plotted as the dashed-dotted line, is reasonably tracked by both our data-driven estimated relationship as well as the monotonically-constrained relationship (dashed line) for values approximately less than 20. The rule-of-thumb bandwidth produces an estimated bid-value relationship that is clearly infeasible.

We see that the use of the rule-of-thumb bandwidth produces a very inaccurate bid-value relationship compared to the true uniform relationship. While this estimator does a poor job of tracking values, it does produce a linear relationship, consistent with the true shape of the bid value relationship. The curve using the data-driven bandwidth does a better job tracking values for low values, but is less reliable for high values, producing a curve that is non-monotonic and swoops away from the true values near the upper end of the plot. Our constrained estimator fixes the pockets of non-monotonicity that appear for the unconstrained estimator and appears to almost exactly mimic the unconstrained curve for low value draws. The allure of the procedures outlined here is that the user has two avenues which to impose smoothness on the underlying bid-value relationship, the bandwidths as well as the constraint weighting to ensure a monotonic relationship. The rule-of-thumb approach precludes this second option.

Panel (b) plots the uniform and constraint weights used in the analysis. We see that for values away from the regions of non-monotonicity, the constraint weights are constant and almost identical to the uniform weights, whereas the weights near the regions of non-monotonicity fluctuate around the uniform weights. A similar pattern is observed in the examples in Hall, Huang, Gifford, and Gijbels (2001).

6   Conclusion

In this paper, we extend a nonparametric method originally proposed for estimating a survival function that can accommodate theoretical restrictions, such as monotonicity, to structural auction settings. The flexibility of the approach permits estimation using data drawn from heterogenous auctions and auctions with reserve prices and differing numbers of bidders. We further extend the method by introducing an automated bandwidth selection process for each of these settings. This is particularly important given that we show that monotonicity, in an empirical auction setting, is directly linked to the bandwidth used

in current nonparametric approaches. This lends further support to the argument that is well-known throughout the statistics and econometrics literature: bandwidth selection is critical, regardless of the setting.

Our work also reveals that errors in bidding by experimental subjects can yield theoretically-inconsistent conclusions. Specifically, applying our estimator, along with our proposed data-driven bandwidth algorithm, to the data collected in Dyer, Kagel, and Levin (1989) yields several small, non-monotonic portions of the estimated equilibrium strategy. At the same time, when using a rule-of-thumb bandwidth, the estimator produces a bid-value relationship which is linear, consistent with the true shape of the bid-value relationship. Unfortunately, this bandwidth also produces a very inaccurate estimate. The primary reason for this difference is that the rule-of-thumb bandwidth assumes that the unknown density is normally distributed.

As stated above, we believe that both the automated bandwidth selection and the ability to impose theoretical constraints on the estimated bidding relationship have value beyond that of structural auctions. Given the importance of monotone comparative statics in economics (Athey 2001, 2002), these techniques should prove indispensable in the application of nonparametric estimation to structural settings.

While we have laid out the framework for constrained nonparametric analysis of auctions, much remains to be done. Future research is needed to extend these methods both within the IPVP, as well as beyond. Within the IPVP, the methods need to be augmented to allow for auction-specific heterogeneity in terms of risk aversion and learning. Outside of the IPVP, these methods can be tailored to the affiliated private value, common value, and conditionally independent private information paradigms that have been developed.

References

[1] Athey, S., 2001. Single crossing properties and the existence of pure strategy equilibria in games of incomplete information. Econometrica 69, 861-890.

[2] Athey, S., 2002. Monotone comparative statics under uncertainty. Quarterly Journal of Economics 108, 187-223.

[3] Athey, S., Haile, P.A., 2002. Identification of standard auction models. Econometrica 70, 2107-2140.

[4] Athey, S., Haile, P.A., 2008. Nonparametric approaches to auctions, in Handbook of Econometrics, Volume 6, edited by J.J. Heckman and E. Leamer. Elsevier: Amsterdam.

[5] Bajari, P., 2001. Comparing competition and collusion: A numerical approach. Economic Theory 18, 187-205.

[6] Bajari, P., Hortaçsu, A., 2005. Are structural estimates of auction models reasonable? Evidence from experimental data. Journal of Political Economy 113, 703-741.

[7] Bajari, P., L. Ye, 2003. Deciding between competition and collusion. Review of Economics and Statistics 85, 971-989.

[8] Beresteanu, A., 2004. Nonparametric estimation of regression functions under restrictions on partial derivatives. Duke University working paper.

[9] Bowman, A. W., P. Hall, T. Prvan, 1998. Bandwidth selction for the smoothing of distribution functions. Biometrika 85, 799-808.

[10] Chak, P.M., N. Madras, B. Smith, 2005. Semi-nonparametric estimation with Bernstein polynomials. Economics Letters 89, 153-156.

[11] Chernozhukov, V., I. Fernandez-Val, and A. Galichon, 2009. Improving point and interval estimates of monotone functions by rearrangement. Biometrika 96(3), 559-575.

[12] Cressie, N.A.C., Read, T.R.C., 1984. Multinomial goodnes-of-fit tests. Journal of the Royal Statistical Society, Series B 46, 440-464.

[13] Du, P., Parmeter, C.F., Racine, J.S., 2010. Constrained Nonparametric Kernel Regression: Estimation and Inference. Working Paper, Department of Statistics, Virginia Tech.

[14] Dyer, D., Kagel, J. H., Levin, D., 1989. Resolving uncertainty about the number of bidders in independent private-value auctions: an experimental analysis. RAND Journal of Economics 20, 268-279.

[15] Gallant, A. R. 1981. On the bias in flexible functional forms and an essentially unbiased form: the Fourier flexible form. Journal of Econometrics 15, 211-245.

[16] Guerre, E., Perrigne, I., Vuong, Q., 2000. Optimal Nonparametric estimation of first-price auctions. Econometrica 68, 525-574.

[17] Hall, P., Huang, L.-S., 2001. Nonparametric kernel regression subject to monotonicity constraints. Annals of Statistics 29, 624-647.

[18] Hall, P., Huang, L.-S., Gifford, J.A., Gijbels, I., 2001. Nonparametric estimation of hazard rate under constraint of monotonicity. Journal of Computational and Graphical Statistics 10, 592-614.

[19] Henderson, D.J., Carroll, R.J., Li, Q., 2008. Nonparametric estimation and testing of fixed effects panel data models. Journal of Econometrics 144, 257-275.

[20] Henderson, D.J., Parmeter, C.F., 2009. Imposing economic constraints in nonparametric regression: Survey, implementation and extension. *Advances in Econometrics: Nonparametric Methods*, Elsevier Science, Volume 25, edited by Q. Li and J. S. Racine, 433-469.

[21] Hotz, V.J., Miller, R.A., 1993. Conditional choice probabilites and the estimation of dynamic models. Review of Economic Studies 60, 497-529.

[22] Imbens, G., Spady, R.H., Johnson, P., 1998. Information theoretic approaches to inference in conditional moment condition models. Econometrica 66, 333-357.

[23] Kneip, A., Simar, L., 1996. A general framework for frontier estimation with panel data. Journal of Productivity Analysis 7, 187-212.

[24] Krasnokutskaya, E., 2009. Identification and estimation in highway procurement auctions under unobserved auction heterogeneity. Review of Economic Studies, forthcoming.

[25] Levitt, S., List, J.A. 2007. What do laboratory experiments measuring social preferences reveal about the real world? Journal of Economic Perspectives 21, 153-174.

[26] Li, Q., Racine, J. 2007. *Nonparametric Econometrics: Theory and Practice*, Princeton University Press.

[27] Li, T., Perrigne, I., Vuong, Q., 2000. Conditionally independent private information in OCS wildcat auctions. Journal of Econometrics 98, 129-161.

[28] Li, T., Perrigne, I., Vuong, Q., 2002. Structural estimation of the affiliated private value auction model. RAND Journal of Economics 33, 171-193.

[29] Li, T., Perrigne, I., Vuong, Q., 2003. Semiparametric estimation of the optimal reserve price in first-price auctions. Journal of Business & Economic Statistics 21(1), 53-64.

[30] McAfee, R. P., D. Vincent, 1992. Updating the reserve price in common-value auctions. American Economic Review, Papers and Proceedings 82, 512-518.

[31] Marron, J.S., Nolan, D., 1989. Canonical kernels for density estimation. Statistics and Probability Letters 7, 195-199.

[32] Martins-Filho, C., Yao, F., 2008. A smooth nonparametric conditional quantile frontier estimator, Journal of Econometrics 143, 317-333.

[33] Matzkin, R., 1994. Restrictions of economic theory in nonparametric methods, *Handbook of Econometrics*, vol. 4 (R.F. Engle and D. L. McFadden eds.) North-Holland: The Netherlands, chapter 42, 2524-2558.

[34] Olley, G. S., Pakes, A., 1996. The dynamics of productivity in the telecommunications equipment industry, *Econometrica* 64(6), 1263-1297.

[35] Owen, A., 1988. Empirical likelihood ratio confidence intervals for a single functional. Annals of Statistics 22, 300-325.

[36] Paarsch, H., 1992. Deciding between the common and private value paradigms in empirical models of auctions. Journal of Econometrics 51, 191-215.

[37] Paarsch, H., Hong, H., 2006. An introduction to the structural econometrics of auction data. MIT Press: Cambridge.

Figure 1: Monotonization for data simulated from a truncated log-normal distribution. Panel (a) represents the 95th percentile of $D_0(\hat{p}) = 1.811407$, the long-dashed line, the GPV estimator, the solid line, as well as the constrained GPV estimator for the same dataset with $\rho = 0.5$, $D_{0.5}(\hat{p}) = 1.846482$ represented by the short-dashed line, and with $\rho = 1$, $D_1(\hat{p}) = 0.003746$ represented by the dotted line. Panel (b) depicts the envelope curves of the values of $\hat{p}$ after the monotonicity constraint had been achieved with $\rho = 0, 0.5$, and 1, again with the respective line types.
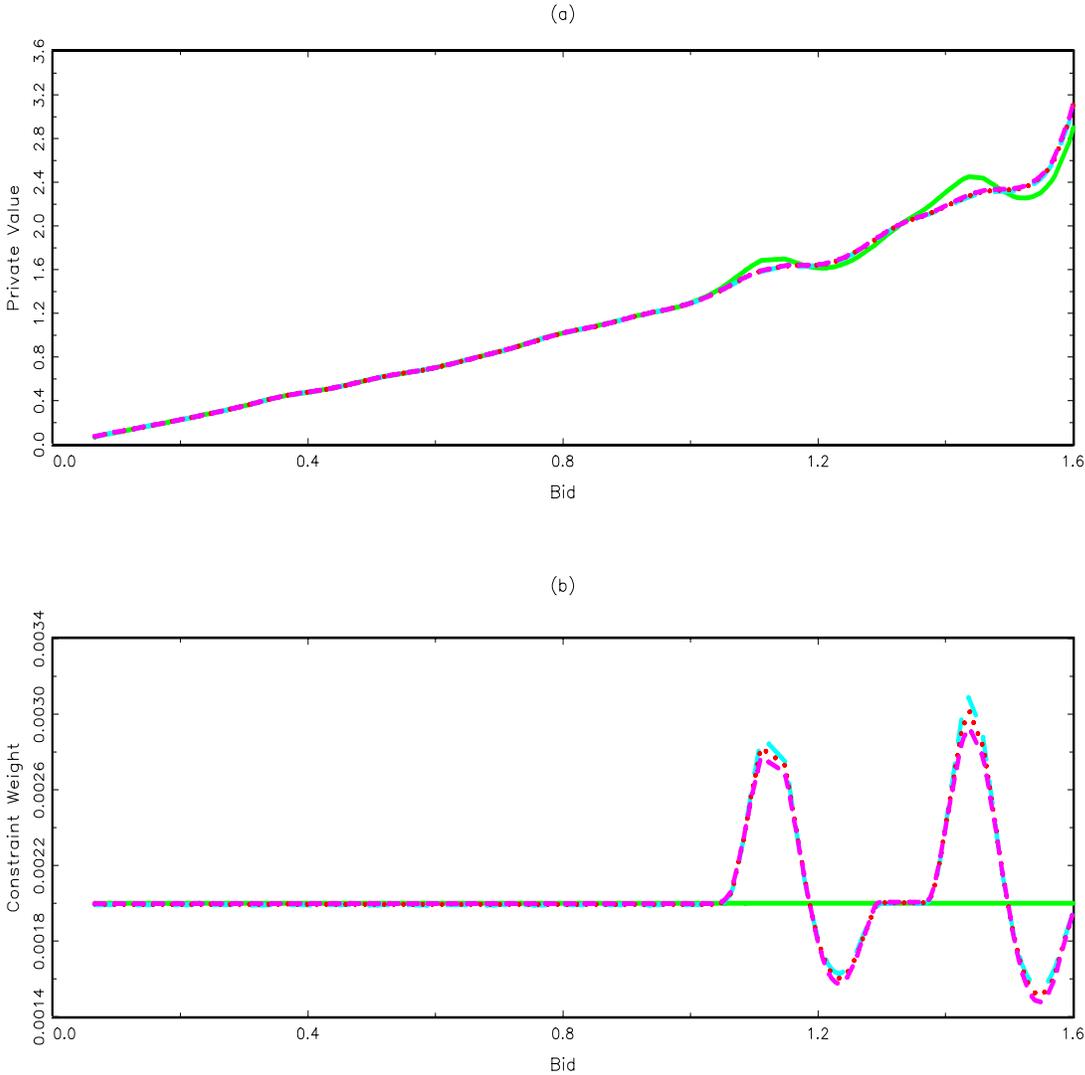
Figure 2: Monotonization for data simulated from a truncated gamma distribution. Panel (a) represents the 95th percentile of $D_0(\hat{p}) = 0.615096$, the long-dashed line, the GPV estimator, the solid line, as well as the constrained GPV estimator for the same dataset with $\rho = 0.5$, $D_{0.5}(\hat{p}) = 0.622522$ represented by the short-dashed line, and with $\rho = 1$, $D_1(\hat{p}) = 0.001259$ represented by the dotted line. Panel (b) depicts the envelope curves of the values of $\hat{p}$ after the monotonicity constraint had been achieved with $\rho = 0, 0.5$, and 1, again with the respective line types.
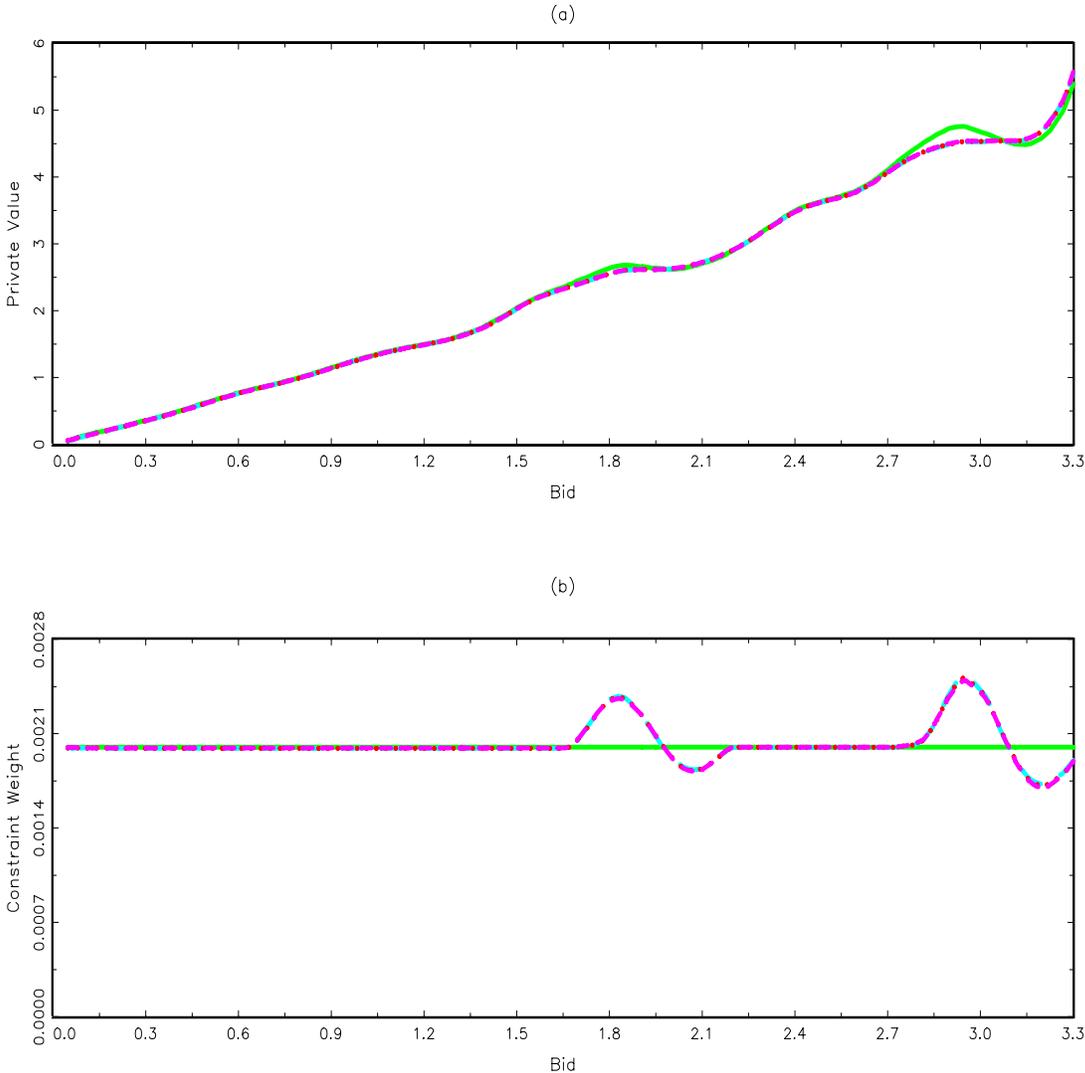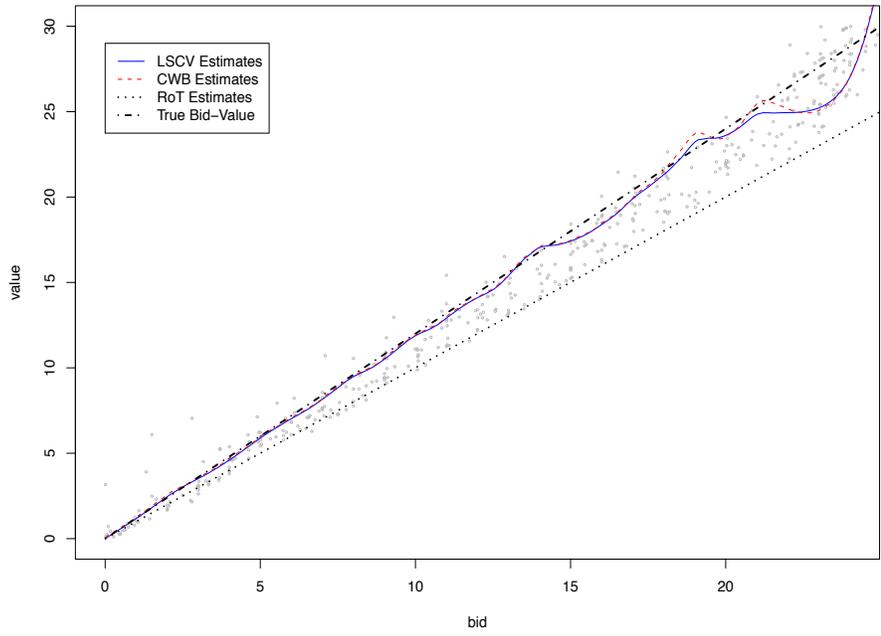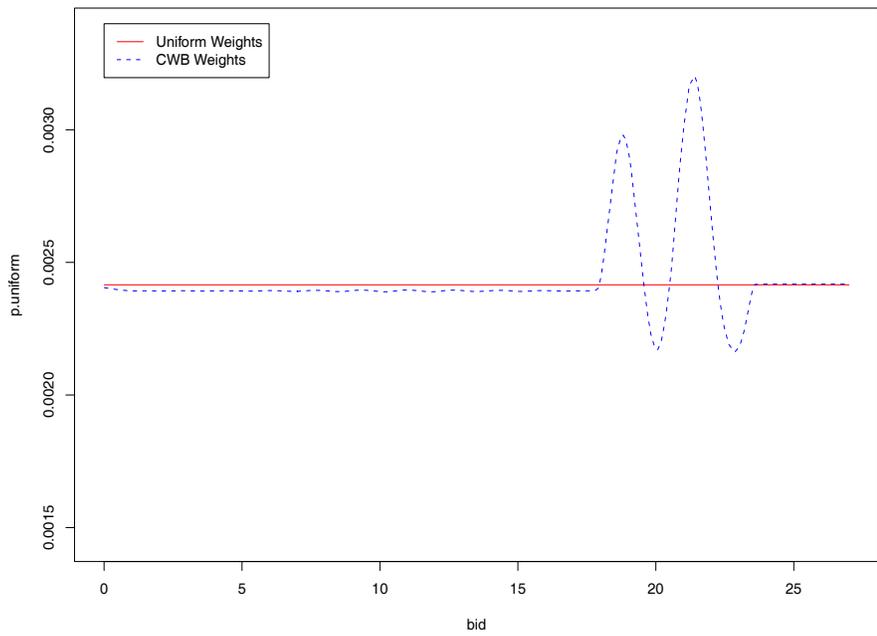
Figure 3: Plots of the Dyer, Kagel, and Levin (1989) first price auction experimental data with six bidders.

(a) Estimated Bid-Value Relationships

(b) Weights