ECON 532: Homework 3

Shabab Ahmed

27 February , 2022

1 Problem 1: Ascending Auctions and SIPV

We know that transaction price is the second highest valuation in this auction format. Let G_W be the distribution of the transaction price (data):

$$G_W(v) = F_{n-1:n}(v) = F(v)^n + nF(v)^{n-1}(1 - F(v)) \qquad (1)$$

We know the number of bidders in our data. We are interested in the distribution F and we have shown that there is an analytic relationship between primitives and the distribution of the data. For the estimation, we just use the $(n-1)$-th order statistic. The identification is based on the fact that the distribution of $i-$th order statistic from n draws is increasing in $F(v)$ and hence invertible. The distribution of i-th order statistic from n draws is given by:

$$F_{i:n}(v) = \frac{n!}{(n-i)!(i-1)!} \int_0^{F(v)} t^{i-1}(1-t)^{n-i}\, dt \qquad (2)$$

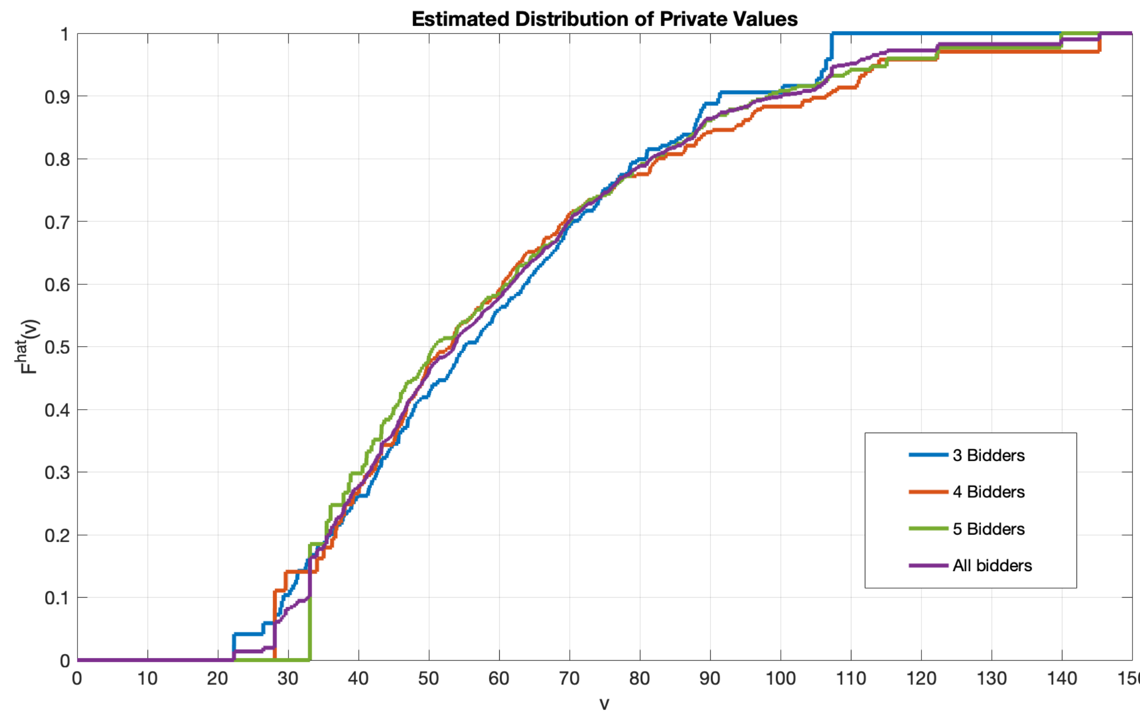
The empirical distribution function $\widehat{F}_{i:n}$ is a pointwise -consistent estimator of $F_{i:n}$. Since the right hand side of (2) is differentiable with respect to $F(v)$, the implicit function theorem and continuous mapping theorem gives us that substituting $\widehat{F}_{i:n}$ for $F_{i:n}$ and solving (2) for $F(v)$ will give a a consistent estimator of $F(v)$. We are interested in $F_{n-1:n}(v)$:

$$\begin{aligned}
F_{n-1:n}(v) &= \frac{n!}{(n-2)!} \int_0^{F(v)} t^{n-2}(1-t)\, dt \\
&= n(n-1) \int_0^{F(v)} t^{n-2} - t^{n-1}\, dt \\
&= n(n-1) \left[\frac{t^{n-1}}{n-1} - \frac{t^n}{n} \right]_0^{F(v)} \\
&= nF(v)^{n-1} + (1-n)F(v)^n
\end{aligned}$$

The above is exactly (1). We have an estimate for $\widehat{G}_W(v) = \widehat{F}_{n-1:n}(v)$ from our data. We substitute that into (1) and rearrange to get the following:

$$\Big(F(v)^n + nF(v)^{n-1}(1-F(v))\Big) - \widehat{G}_W(v) = \Big(F(v)^n + nF(v)^{n-1}(1-F(v))\Big) - \widehat{F}_{n-1:n}(v) = 0 \qquad (3)$$

We use MATLAB's *roots* function to solve equation (3). We have to exercise caution in doing so as the function can return implausible roots for our purpose. We only considered the roots that are between $[0, 1]$ and we made sure to count one instance of a repeated root. For some roots, there was the issue of getting a complex number of the form $(a + 0i)$ and we just considered the real part of those roots. The following figure contains the estimated CDF of the bidders' private values:



I calculated the distribution for all bidders by using the average of the 3 bidders, 4 bidders and 5 bidders cases. The code is included in the zip file *ShababAhmed_code.zip*. The relevant files are *Problem1.m*, *pricedist.m* and *valuedist.m*.

2 Problem 2: Haile and Tamer

Let Δ be the minimum bid increment. Let $G_{i:n}(v) = P\left(b_{i:n} \leq v\right)$. We have the following:

Upper Bound: Assume that bidders never bid more than their valuations.

- Estimate probability that i-th highest bid with n bidders is below v:

$$\widehat{G}_{i:n}(v) = \frac{1}{T_n} \sum_{t=1}^{T} \mathbb{1}\{n_t = n, b_{i:n_t} \leq v\} \xrightarrow{p} G_{i:n}(v)$$

- Lemma 1: $b_{i:n} \leq v_{i:n}$ for all $i \leq n$
- Lemma 1 implies

$$F_{i:n}(v) \leq G_{i:n}(v) \text{ for all } i$$

Lower Bound:

- Lemma 3: $v_{n-1:n} \leq b_{n:n} + \Delta$
- From Lemma 3:
$$F_{n-1:n}(v) \geq G^{\Delta}_{n:n}(v)$$

where $G_{n:n}(v) = P\left(b_{n:n} + \Delta \leq v\right)$.Then, using a monotone operator ϕ we have that:

$$F(v) \leq F_U(v) = \min_{n,i} \phi\left(G_{i:n}(v), i, n\right)$$
$$F(v) \geq F_L(v) = \max_{n} \phi\left(G^{\Delta}_{n:n}(v), n-1, n\right)$$

where

$$F_{i:n}(v) = \frac{n!}{(n-i)!(i-1)!} \int_0^{F(v)} t^{i-1}(1-t)^{n-i}\, dt$$
$$\implies F(v) = \phi\left(F_{i:n}(v), i, n\right)$$

Haile and Tamer (2003) deals with an English auction which is a first price auction. Therefore, the transaction price is the highest bid and so we have data on $b_{n:n}$ allowing us to estimate $G_{n:n}(v)$ with $\widehat{G}_{n:n}(v)$.

Notice:

$$
\begin{aligned}
F_{n:n}(v) &= n \int_0^{F(v)} t^{n-1} \\
&= n \left[\frac{t^n}{n} \right]_0^{F(v)} \\
&= F(v)^n \\
\implies F(v) &= F_{n:n}(v)^{\frac{1}{n}} = \phi(F_{n:n}(v))
\end{aligned}
$$

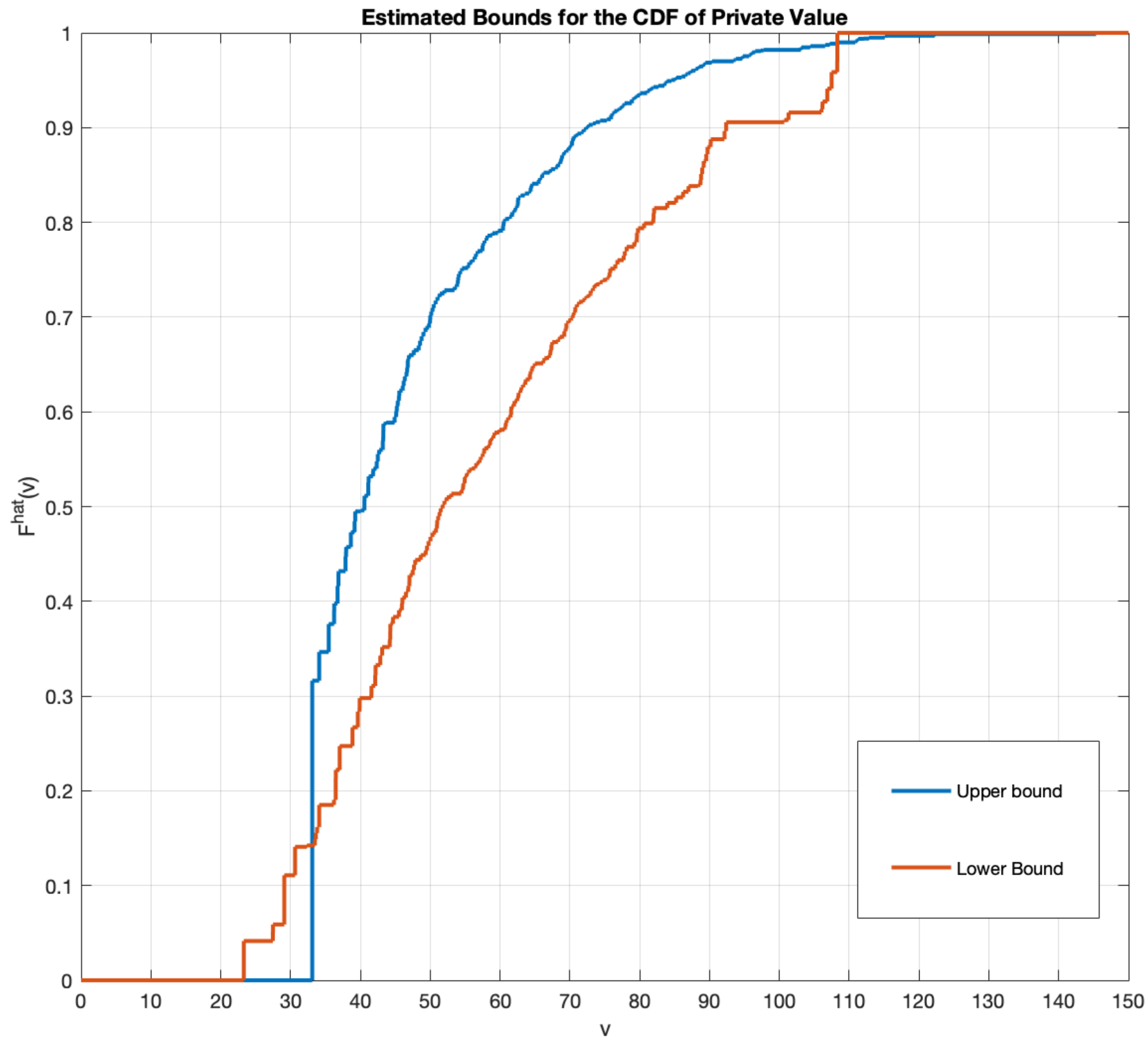
Thus, for the upper bound we have the following:

$$
F(v) \leq F_U(v) = \min_n \phi\left(\widehat{G}_{n:n}(v)\right) = \min_n \left(\widehat{G}_{n:n}(v)\right)^{\frac{1}{n}}
$$

For the lower bound, we can easily obtain $\widehat{G}^{\Delta}_{n:n}(v)$ an estimate of $G^{\Delta}_{n:n}(v)$ because we observe $b_{n:n}$ and also know delta. Hence:

$$
F(v) \geq F_L(v) = \max_n \phi\left(\widehat{G}^{\Delta}_{n:n}(v)\right)
$$

where ϕ is the inverse of the right hand side of (1) in Problem 1 because here we are concerned with $F_{n-1:n}$. We will again use the approach in Problem 1 where we simply find the roots of (3) in Problem 1 instead of computing the inverse. One thing to note is that for the calculation of $\widehat{G}_{n:n}(v)$ and $\widehat{G}^{\Delta}_{n:n}(v)$ we can use the same function we used in Problem 1 to calculate transaction prices. This is because we are still calculating the empirical CDF of the prices/highest bids. We just have to plug in the right value of n and also add an argument for Δ. We will have $\Delta = 0$ for Problem 1 and the upper bound. We will have $\Delta = 1$ for the calculation of the lower bound. The following figure provides the Haile and Tamer (2003) estimated bounds on the CDF of the bidders' private values:



We can see that the bounds cross. This is understandable as the bounds may be noisy due to the finite sample. We know from Haile and Tamer (2003) that bounds may cross in finite sample. The code is included in the zip file *ShababAhmed_code.zip*. The relevant files are *Problem2.m*, *pricedist.m*, *calculatebounds.m* and *valuedist.m*.

---

3 Problem 3:

3.1 First-Price Sealed Bid Auctions with SAPV

We have an APV model here. We follow the approach of Li, Perrigne and Vuong (2002) for this problem. We will denote the conditional distribution of B_1 given b_1 by $G_{B_1|b_1}(\cdot|\cdot)$ and its density by $g_{B_1|b_1}(\cdot|\cdot)$. LPV(2002) shows that we can express each private value as a function of its corresponding bid as well as the distribution and density of observed bids:

$$v = b + \frac{G_{B_1|b_1}(b|b)}{g_{B_1|b_1}(b|b)} \qquad (4)$$

We can estimate $G_{B_1|b_1}(\cdot|\cdot)$ and $g_{B_1|b_1}(\cdot|\cdot)$ from observed bids, then we can use (4) to recover the private values for all bidders. Extending GPV (2000), the paper uses the following two-step estimation procedure:

Step 1. Construct a sample of pseudo-private values based on (4) using nonparametric estimates of $G_{B_1|b_1}(\cdot|\cdot)$ and $G_{B_1|b_1}(\cdot|\cdot)$ from observed bids.

Step 2. Use the pseudo-sample constructed in Step 1 to estimate nonparametrically the joint density of bidders' private values.

We have $n = 4$ in our sample. Let L be the number of auctions. In the first step, using the observed bids $\{b_{il}; i = 1, \dots, n, l = 1, \dots, L\}$ we estimate nonparametrically the ratio $\frac{\tilde{G}_{B_1|b_1}(\cdot|\cdot)}{\tilde{G}_{B_1|b_1}(\cdot|\cdot)}$ by $\frac{\widehat{G}_{B_1:b_1}(\cdot,\cdot)}{\widehat{g}_{B_1:b_1}(\cdot,\cdot)}$ where

$$\widehat{G}_{B_1,b_1}(B,b) = \frac{1}{Lh_G} \sum_{l=1}^{L} \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(B_{il} \leq B) K_G\left(\frac{b - b_{il}}{h_G}\right)$$
$$\widehat{g}_{B_1,b_1}(B,b) = \frac{1}{nLh_g^2} \sum_{l=1}^{L} \sum_{i=1}^{n} K_g\left(\frac{B - B_{il}}{h_g}, \frac{b - b_{il}}{h_g}\right)$$

where h_G and h_g are some bandwidths, K_G and K_g are kernels. We will use the triweight kernel which is defined in Part b) of this problem for K_g. We will use the product of two triweight kernels for K_g. We will use $h_G = 2.978 \times 1.06\widehat{\sigma}_b \times (nL)^{-\frac{1}{5}}$ where $\widehat{\sigma}_b$ is the empirical standard deviation of the bids. We will use $h_g = 2.978 \times 1.06\widehat{\sigma}_b \times (nL)^{-\frac{1}{6}}$. This is not quite in line with Proposition 2 of the paper but if we increase the denominator of the exponent the bounds become extremely narrow. We then calculate estimates of the unobserved private values v_{il} as follows:

$$\widehat{v}_{il} = b_{il} + \frac{\widehat{G}_{B_1,b_1}(b_{il}, b_{il})}{\widehat{g}_{B_1,b_1}(b_{il}, b_{il})}$$

The paper suggests to do some trimming. Let B_{max} be the maximum of the observed bids. The paper removes bids that are outside the interval $[h_g, B_{max} - h_g]$. However, we do not do trimming in this paper because it makes the number of observations too small using their recommended trimming. The interval $[h_g, B_{max} - h_g]$ becomes too narrow. In the second step, we use the pseudo-sample to estimate nonparametrically the joint density

$$\widehat{f}(v_1, \dots, v_n) = \frac{1}{Lh_f^n} \sum_{l=1}^{L} K_f\left(\frac{v_1 - \widehat{v}_{1l}}{h_f}, \dots, \frac{v_n - \widehat{v}_{nl}}{h_f}\right)$$

where h_f is a bandwidth and K_f is the product of n trighweight kernels. We let $h_f = 2.978 \times 1.06\widehat{\sigma}_b \times (4!L)^{-\frac{1}{17}}$. We do not implement this in the code but it would be pretty simple to implement because all we need is the triweight kernel function. This has already been defined in our code. We implement the joint CDF function instead.

We know that bids are strictly increasing functions of the valuations, so that quantiles of the bid distribution map through the inverse equilibrium bid function to the same quantiles of the value distribution. Therefore, we can calculate the 25^{th} and 75^{th} percentiles of the bids and use the function $v(b)$ to get the corresponding percentiles of the values.The following is a table containing the values along with the corresponding percentiles:

Percentile	u_1	u_2	u_3	u_4
[25, 25,25,25]	49.1504	38.0483	42.5626	40.6480
[25, 25, 25, 75]	49.1504	38.0483	42.5626	156.3653
[25, 25, 75, 25]	49.1504	38.0483	146.3628	40.6480
[25, 25, 75, 75]	49.1504	38.0483	146.3628	156.3653
[25, 75, 25, 25]	49.1504	145.0387	42.5626	40.6480
[25, 75, 25, 75]	49.1504	145.0387	42.5626	156.3653
[25, 75, 75, 25]	49.1504	145.0387	146.3628	40.6480
[25, 75, 75, 75]	49.1504	145.0387	146.3628	156.3653
[75, 25, 25, 25]	153.1781	38.0483	42.5626	40.6480
[75, 25, 25, 75]	153.1781	38.0483	42.5626	156.3653
[75, 25, 75, 25]	153.1781	38.0483	146.3628	40.6480
[75, 25, 75, 75]	153.1781	38.0483	146.3628	156.3653
[75, 75, 25, 25]	153.1781	145.0387	42.5626	40.6480
[75, 75, 25, 75]	153.1781	145.0387	42.5626	156.3653
[75,75,75,25]	153.1781	145.0387	146.3628	40.6480
[75,75,75,75]	153.1781	145.0387	146.3628	156.3653

The following table contains $F_U(u_1, u_2, u_3, u_4)$. We also include a column for $F_I(u_1, u_2, u_3, u_4)$ which gives us distribution under independence assumption. We know that $F_I(u_1, u_2, u_3, u_4) = \prod_{i=1}^{4} F_U(u_i)$.

Percentile	F_U	F_I
[25, 25,25,25]	0	0.0039
[25, 25, 25, 75]	0.0080	0.0117
[25, 25, 75, 25]	0.0040	0.0117
[25, 25, 75, 75]	0.0220	0.0352
[25, 75, 25, 25]	0.0060	0.0117
[25, 75, 25, 75]	0.0300	0.0352
[25, 75, 75, 25]	0.0280	0.0352
[25, 75, 75, 75]	0.1040	0.1055
[75, 25, 25, 25]	0.0180	0.0117
[75, 25, 25, 75]	0.0480	0.0352
[75, 25, 75, 25]	0.0320	0.0352
[75, 25, 75, 75]	0.0980	0.1055
[75, 75, 25, 25]	0.0380	0.0352
[75, 75, 25, 75]	0.1080	0.1055
[75,75,75,25]	0.1020	0.1055
[75,75,75,75]	0.3140	0.3164

It is hard to tell from this table alone whether the symmetry or independence assumption holds. We would have to do other tests to check either of the assumptions. We can see that F_U and F_I are different but the values are pretty close. So, it is not exactly clear whether independence holds. If symmetry assumption is true, we should expect that any permutation of, for example, $[25, 25, 75, 75]$, should yield the same value. This means that we should expect $[75, 75, 25, 25]$, for example, should give the same F_U. We can see that this is not necessarily true. Again, the values are close by but they are not exactly equal. To further hone in on symmetry, we provide summary statistics for the bidders values computed using the method in the paper.

	Bidder 1	Bidder 2	Bidder 3	Bidder 4
Mean	107.1608	102.2801	102.1176	104.1725
25^{th} percentile	49.1504	38.0483	42.5107	40.6480
50^{th} percentile	103.3926	96.3161	99.6495	99.5515
75^{th} percentile	153.2705	145.1327	146.2634	156.5234
Min	0.4514	0.0075	0.1897	0.1937
Max	277.9445	279.5133	275.9452	279.8161

We can see that there are some disparities in the distribution. In particular, bidder 1's mean and median seem to differ considerably from the other bidders. Bidders 2 and 4 seem to be closer to each other, thus, it is likely that the symmetry assumption is true for them. However, we cannot conclude the same about all the bidders.

The code is included in the zip file *ShababAhmed_code.zip*. The relevant files are *Problem3a.m*, *pseudovalue_APV.m*, *triweight_APV.m*, *calculateG_APV.m*, *calculategtild_APV.m*, *jointcdf.m* and *quantilevals.m*.

3.2 First-Price Sealed Bid Auctions with SIPV

We can use the GPV (2000) approach for this problem. We can recover bidder i's value as a function of their bid:

$$v_i = b_i^* + \frac{G(b_i^*)}{(I-1)g(b_i^*)}$$

where I is the number of bidders and $G(b) = P(\beta(v) \leq b)$. $G(b_i)$ and $g(b_i)$ can be easily estimated by the following:

$$\tilde{G}(b) = \frac{1}{IL}\sum_{l=1}^{L}\sum_{p=1}^{I}\mathbb{1}\left(B_{lp} \leq b\right)$$
$$\tilde{g}(b) = \frac{1}{ILh_g}\sum_{l=1}^{L}\sum_{p=1}^{I}K_g\left(\frac{b - B_{lp}}{h_g}\right)$$

where L is the number of auctions, where h_g is a bandwidth and $K_g(\cdot)$ is a kernel with a compact support. We trim the data following the steps in GPV (2000). Let p_g be the length of the support of K_g, B_{min} be the minimum bid observed in the data and B_{max} be the maximum bid observed in the data. We trim the data by leaving out bids in the dataside which lie outside the following interval: $[B_{min} + \frac{\rho_g h_g}{2}, B_{max} - \frac{\rho_g h_g}{2}]$.

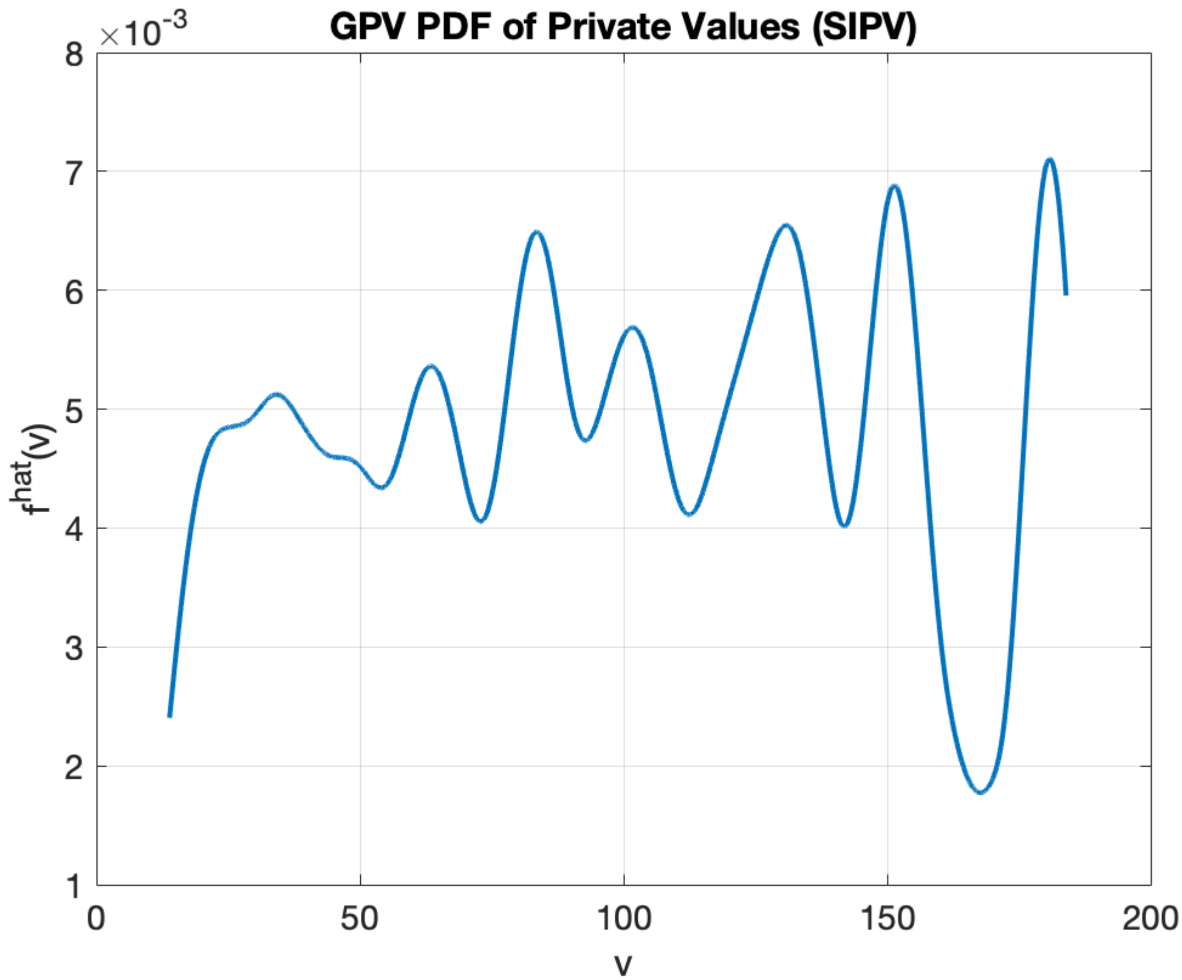
We use the triweight kernel following GPV (2000): $\frac{35}{32}(1-u^2)^3\mathbb{1}(|u| \leq 1)$. This implies that $_g = 2$.We also use $h_g = 1.06 * \widehat{\sigma}_b * (IL)^{-\frac{1}{5}}$ where $\widehat{\sigma}_b$ is the estimated standard deviations of the observed bids. Once, we have calculated the $\tilde{G}(b)$ and $\tilde{g}(b)$ using the observed bids, the triweight kernel function and the bandwidth we can calculate pseudo private values $\widehat{V}$:

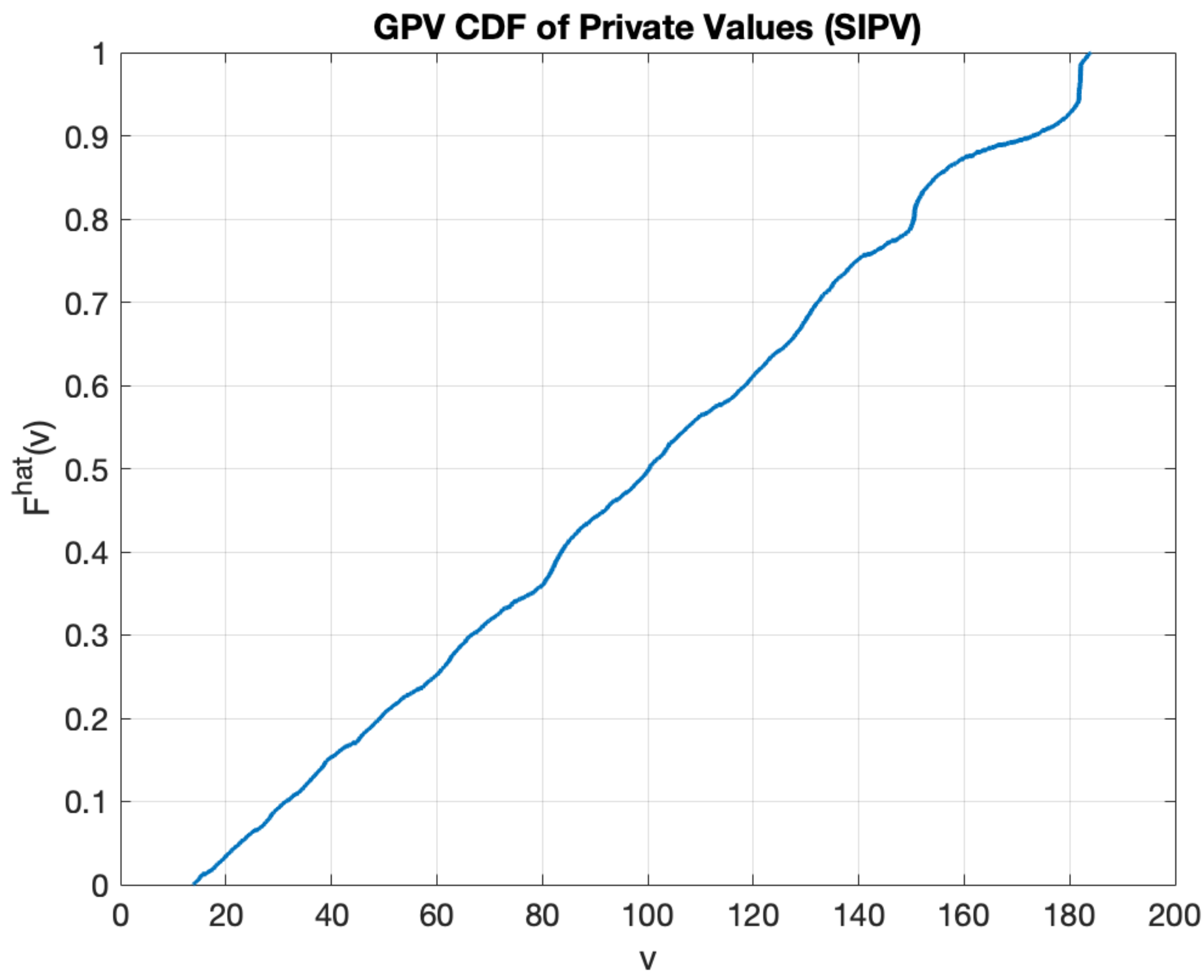
$$\widehat{V}_{il} = b_{il} + \frac{1}{I-1}\frac{\tilde{G}\left(b_{il}\right)}{\tilde{g}\left(b_{il}\right)}$$

We also include an estimate of the pdf of private values using the following:

$$\widehat{f}(v) = \frac{1}{ILh_f}\sum_{l=1}^{L}\sum_{i=1}^{I}K_f\left(\frac{v - \widehat{V}_{il}}{h_f}\right)$$

We follow GPV (2000) again. We use the triweight kernel for the K_f and we let $h_f = 1.06\widehat{\sigma} - v(IL_T)^{\frac{-1}{5}}$ where $\widehat{\sigma}_v$ is the estimated standard deviations of the trimmed pseudo private values $\widehat{V}$ and L_T is the number of auctions remaining after the trimming. We also provide a graph of the estimated CDF of the private values. We use the pseudo sample to construct the empirical cdf. The estimated PDF and CDFs of the pseudo private values are provided below:


GPV PDF of Private Values (SIPV)
×10-3
8
7
6
5
4
3
2
1
fhat(v)
0
50
100
150
200
v



The code is included in the zip file *ShababAhmed_code.zip*. The relevant files are *Problem3b.m*, *pseudovalue.m*, *trimming.m*, *triweight.m*, *calculateG.m*, *calculategtild.m*, *GPVpdf.m* and *GPVcdf.m*.
