# 1  Estimating Value Distribution from Equilibrium Bids

We will look at the case of a single-item First Price Auction. Our goal will be to estimate the distribution of valuations of the players, by observing samples of bid vectors of players from multiple first price auctions.

In particular, we consider the case where we have a data set of $m$ First Price Auctions, each among $n$ bidders. The value $v_i^t$ of each bidder $i$ in each first price auction $t$, is drawn i.i.d. from the same distribution with CDF $F(\cdot)$ and density $f(\cdot)$ and support $[0, \bar{v}]$.

We will not observe these values $v_i^t$, but we will instead be observing the bid of each player $b_i^t$. Moreover, we will be assuming that player's play according to a symmetric Bayes-Nash equilibrium that corresponds to the value distribution $F$. In other words, the players know the distribution $F$ and they have computed a symmetric Bayes-Nash equilibrium of a single item first price auction for that distribution. Such a Bayes-Nash equilibrium (BNE) is mapping from a value $v \in [0, \bar{v}]$ to a bid $b(v)$, such that for all $b_i' \in [0, \bar{v}]$:

$$\mathbb{E}_{v_{-i}} \left[ u_i(b(v_1), \ldots, b(v_n); v_i) \right] \geq \mathbb{E}_{v_{-i}} \left[ u_i(b_i', b_{-i}(v_{-i}); v_i) \right] \tag{1}$$

Thus what we observe is $b_i^t = b(v_i^t)$ for each player $i$ and auction $t$. Our goal is to compute estimates $\hat{F}$ and $\hat{f}$ of $F$ and $f$ from observations of such bid vectors.

# 2  Identification: Infinite Data Limit

We first show that if we had an infinite amount of data, then we can reverse engineer and compute the distribution of values from simply observing the distribution of bids. This task is called the identification problem, where we ignore sampling errors in the data and simply look at the infinite data limit, typically referred to as the *population limit*. This problem is sort of a sanity check, that at least if you didn't have a sampling error, the problem of computing the quantity of interest (here the value distribution) is feasible.

To address this problem we will make some regularity assumptions about our distribution of values:

**Assumption 1.** *The CDF function $F$ is strictly increasing, continuous and differentiable in the support $[0, \bar{v}]$. Hence, the density is always lower bounded by some strictly positive constant in the support.*

We first refer to a recent result of [4] that shows that in a single item first price auction with a random tie-breaking rule, there only exist symmetric Bayes-Nash equilibria, and moreover there is only a unique Bayes-Nash equilibrium, under which the bid of a player is a strictly increasing function of his value.

**Lemma 1** ([4]). *Under assumption 1, the equilibrium of the first price auction is unique, symmetric and $b(\cdot)$ is strictly increasing, continuous and differentiable. Moreover, $b(0) = 0$ and $b(\bar{v}) = \bar{b}$ for some $\bar{b} \leq \bar{v}$, and $b(v)$ is defined by Myerson's payment identity:*

$$b(v) = v - \frac{1}{F(v)^{n-1}} \int_0^v F(z)^{n-1} dz \tag{2}$$

Given this basic theorem, we will now try to understand the relation between bids and values, by analyzing the first order conditions of a symmetric BNE. Let $b(\cdot)$ be the BNE bidding function. Moreover, let $G(\cdot)$ be the CDF of the bid distribution generated by drawing a value $v$ from $F$ and then mapping it to $b(v)$. Also let $g(\cdot)$ be the density of $G$. Observe that by the properties of $b(\cdot)$, $G$ is continuous and admits a density.

Since a player is playing according to an equilibrium we can connect his bid with his value by the best-response condition. If a player submits a bid $b$, his probability of winning is the probability that all other players submit a value below $b$. The latter happens with probability $G(b)^{n-1}$. Thus his expected utility is:

$$u(b; v) = (v - b) \cdot G(b)^{n-1} \tag{3}$$

Since a player's bid $b(v)$ maximizes his utility, it must be that the derivative of the above expression is zero at $b = b(v)$. The derivative of the latter with respect to $b$ is:

$$\frac{\partial u(b; v)}{\partial b} = -G(b)^{n-1} + (v - b)(n - 1)G(b)^{n-2}g(b) \tag{4}$$

Since it must be equal to zero at $b = b(v)$, we get:

$$v - b(v) = \frac{G(b(v))}{(n - 1)g(b(v))} \tag{5}$$

Alternatively, if we denote with $\xi(\cdot) = b^{-1}(\cdot)$, then by doing a change of variables in the above equation, we have:

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

Moreover, since $b(\cdot)$ is strictly increasing, continuous and differentiable, we have that $\xi(\cdot)$ is also strictly increasing, continuous and differentiable.

Thus if we knew the distribution of bids, i.e. $G(\cdot)$ and $g(\cdot)$, then given the bid of a player $b$ we can reverse engineer his value $\xi(b)$. Thus we can conclude, that we can *identify* the distribution of values, since:

$$F(v) = \Pr[V \le v] = \Pr[b(V) \le b(v)] = G(b(v)) = G(\xi^{-1}(v)) \tag{7}$$

In other words, if we want to estimate the probability mass that falls below a number $v$, we can do it by first computing the equilibrium bid $\xi^{-1}(v)$, that corresponds to this value. This is computable when one has access to $G$ and $g$, by Equation (6). Then we compute how much mass does $G$ put below $\xi^{-1}(v)$. The latter is also computable if we are given access to $G$.

The key property that enabled us to do the latter identification, was that the best response condition of a player only depends on his underlying value and on the behavior of other bidders, where the latter behavior is also observable to the econometrician. This enables us to estimate the value of the player, given access to the behavior of other players. This is a property that is useful in estimation in many other games of incomplete information and not only auctions (see for instance [2] for a survey on identification and estimation in games of incomplete information).

## 3 Estimation: Finite Data

We now turn to the finite data case and show how we can estimate the distribution of values. The identification strategy of the previous section reveals an easy two-stage approach for estimating the distribution of values. In what we do below we don't really use the auction from which the bid of a player came from, just the fact that it was an equilibrium bid from an $n$-bidder auction. Thus for simplicity we assume we have access to $m$ samples of bids $b_1, \ldots, b_m$ drawn i.i.d. from the equilibrium bid distribution $G$, and we want to estimate $F$ and $f$.

If we knew the true $G$ and $g$, then given samples $b^t$ of bids of each player, we can compute samples $v^t$ of values, drawn from $F$ (using Equation (6)). Then we would be back to the simple setting of estimating the CDF and PDF of a distribution when having access to i.i.d. samples from that distribution. Since

we have i.i.d. samples from $G$ we can estimate $G$ and $g$ up to a reasonable accuracy. Thus we can compute values $\hat{v}^t$ that are almost equal to $v^t$, by using an estimated version of Equation (6). Hence, this leads to the following two-stage strategy:

1. (First stage.) Compute estimates $\hat{G}$ and $\hat{g}$ of the distribution of bids.

2. (Second stage.) Invert each bid to compute the corresponding value, from an estimated analogue of Equation (6), i.e. :

$$\hat{v}_t = \hat{\xi}(b_t) \equiv b_t + \frac{\hat{G}(b_t)}{(n-1)\hat{g}(b_t)} \tag{8}$$

Estimate $\hat{F}$ and $\hat{f}$, based on values $\hat{v}_t$, treating them as if they were the true values, hence, i.i.d. samples from $F$.

For the estimation of CDFs, $\hat{G}$ and $\hat{F}$, we will use the empirical CDF, i.e.:

$$\hat{G}(b) = \frac{1}{m} \sum_{t=1}^{m} 1\{b_t \leq b\} \tag{9}$$

$$\hat{F}(v) = \frac{1}{m} \sum_{t=1}^{m} 1\{\hat{v}_t \leq v\} \tag{10}$$

For the estimation of the densities $\hat{g}$ and $\hat{f}$ we will use Kernel density estimation:

$$\hat{g}(b) = \frac{1}{m} \sum_{t=1}^{m} \frac{1}{h_g} K_g \left( \frac{b_t - b}{h_g} \right) \tag{11}$$

$$\hat{f}(b) = \frac{1}{m} \sum_{t=1}^{m} \frac{1}{h_f} K_f \left( \frac{\hat{v}_t - v}{h_f} \right) \tag{12}$$

where $K_f$ and $K_g$ is some Kernel function. This could be for instance the uniform Kernel: $K_u(z) = \frac{1}{2}1\{|z| \leq 1\}$ or the Epanechnikov Kernel $K_e(z) = \frac{3}{4}(1-z^2)1\{|z| \leq 1\}$. Since we will be using Kernel density estimation we will need to make some Lipschitzness assumptions about our densities. In particular, we will assume:

**Assumption 2.** *The density of the value distribution $f(\cdot)$ is $\lambda$-Lipschitz.*

For the remainder of the section we will only be using the very simple uniform Kernel and give finite sample uniform error bounds of our estimate $\hat{f}$ of the PDF of values. Similar analysis can give bounds on the estimate $\hat{F}$, which is an easier case. Our goal will be to show the following guarantee: [1]

**Theorem 1.** *Under assumptions 1 and 2, and with $K_f, K_g$ being the uniform Kernel, and $h_g = O(m^{-1/4})$ and $h_f = O(m^{-1/8})$, for any fixed interior sub-interval $C_v(\epsilon) = [\epsilon, \bar{v} - \epsilon]$ of the support of $F$, we have: if $m$ is larger than some constant, with probability $1 - \delta$*

$$\sup_{v \in C_v(\epsilon)} |\hat{f}(v) - f(v)| \leq O\left( \left( \frac{\log(1/\delta)}{m} \right)^{1/8} \right) \tag{13}$$

## 3.1 Analyzing First Stage Errors

We start by analyzing the uniform errors of our first stage estimates, $\hat{G}$ and $\hat{g}$. Since we are using the empirical CDF for $\hat{G}$, by the DKW in equality that we show in Lecture 17 we have:

---

[1]If one makes the stronger assumption that $f(\cdot)$ has $R+1$ continuous bounded derivatives, then a faster convergence rate of $O\left( \left( \frac{\log(m)}{m} \right)^{R/(2R+3)} \right)$, can be achieved, via a tighter analysis (see [5]). For instance, for $R = 1$ (which is a slightly stronger condition than our Lipschitzness condition), the latter gives a rate with $1/5$ in the exponent rather than $1/8$.

**Lemma 2.** *With probability $1 - \delta$:*

$$\sup_{b \in B} |\hat{G}(b) - G(b)| \leq O\left(\sqrt{\frac{\log(1/\delta)}{m}}\right) \tag{14}$$

Thus it remains to show the properties of $\hat{g}$. We start by stating that under assumptions 1 and 2, the PDF $\hat{g}$ must also be Lipschitz continuous and must also be bounded away from 0 in any fixed interior subset of the support of distribution $G$.

**Lemma 3.** *Under Assumptions 1 and 2, the CDF $G$ is $\lambda_G$-Lipschitz in $[0, \bar{b}]$ and the density $g$ is bounded away from 0 in its support and is $\lambda_g$-Lipschitz for any fixed interior subset $C_b(\epsilon) = [\epsilon, \bar{b} - \epsilon]$ of its support.*

**Proof.** We give a sketch of the proof. Since the equilibrium is symmetric and strictly increasing, we know that the expected allocation of a player with value $v$ is $x(v) = F(v)^{n-1}$. Thus by Myerson's payment identity:

$$b(v) \cdot x(v) = p(v) = v \cdot x(v) - \int_0^v x(z)dz \Leftrightarrow b(v) = v - \frac{1}{F(v)^{n-1}} \int_0^v F(z)^{n-1}dz \tag{15}$$

Observe that since $F$ is differentiable $b(\cdot)$ is also differentiable. Hence, its inverse $\xi(\cdot)$ is also differentiable and $\xi'(b) = \frac{1}{b'(\xi(b))}$. Moreover, since it is strictly increasing, $b'(v) \geq \delta$, hence, it's inverse is Lipschitz (since $\xi'(b) \leq 1/\delta$). Since $G(b) = F(\xi(b))$, and both $F$ and $\xi$ are Lipschitz, $G$ is also Lipschitz. Moreover, since $F$ and $\xi$ are both strictly increasing we also get that $G$ is strictly increasing, and hence $g$ is bounded away from 0. Now by re-arranging equation (6), we can write: $g(b) = \frac{G(b)}{(n-1)(\xi(b)-b)}$. Since $G$ is Lipschitz, $\xi$ is Lipschitz and $\xi(b) - b = \frac{1}{F(\xi(b))^{n-1}} \int_0^{\xi(b)} F(z)^{n-1}dz$ is bounded away from zero in $C_b(\epsilon)$, we also get that $g$ is Lipschitz. $\square$

We are now ready to show that for any interior of the support $[0, \bar{b}]$, the Kernel density estimator, with the uniform Kernel has small error.

**Lemma 4.** *For any interior $C_b(2\epsilon)$, and for the uniform Kernel, with bandwidth $h_g \leq \epsilon$, we have that: with probability $1 - \delta$*

$$\sup_{b \in C_b(2\epsilon)} |\hat{g}(b) - g(b)| \leq \lambda_g h_g + \frac{1}{h_g} \frac{O(\sqrt{\log(1/\delta)})}{\sqrt{m}} \tag{16}$$

*Hence, for $h_g = O(m^{-1/4})$ and for $m \geq 1/\epsilon^4$, we have that the latter bound is at most $O\left(\left(\frac{\log(1/\delta)}{m}\right)^{1/4}\right)$.*

**Proof.** Since the Kernel density estimator at any point in $C_b(2\epsilon)$ uses only estimates from a region that is in $C_b(\epsilon)$ and since $g$ is $\lambda_g$-Lipschitz in $C_b(\epsilon)$, we get the Theorem by the analysis of the uniform Kernel for Lipschitz densities that we did in Lecture 19. $\square$

## 3.2 Value Inversion and Inversion Error

Based on the first stage errors, we can now upper bound the error in our second stage estimation of the value of each bidder, i.e. bounding $|\hat{v}_t - v_t|$. We first provide a uniform error bound between $\xi$ and $\hat{\xi}$.

**Lemma 5.** *For any fixed interior $C_b(2\epsilon)$, for $\epsilon > 0$, and for $m$ larger than some constant, we have that:*

$$\sup_{b \in C_b(2\epsilon)} |\hat{\xi}(b) - \xi(b)| \leq O\left(\left(\frac{\log(1/\delta)}{m}\right)^{1/4}\right) \tag{17}$$

**Proof.** By expanding $\hat{\xi}$ and $\xi$, we have that for any $v \in C_b(2\epsilon)$:

$$
\begin{aligned}
|\hat{\xi}(b) - \xi(b)| &= \left| b + \frac{\hat{G}(b)}{(n-1)\hat{g}(b)} - b - \frac{G(b)}{(n-1)g(b)} \right| = \left| \frac{\hat{G}(b)}{(n-1)\hat{g}(b)} - \frac{G(b)}{(n-1)g(b)} \right| \\
&= \frac{1}{\hat{g}(b) \cdot g(b)} \left| \hat{G}(b)g(b) - G(b)\hat{g}(b) \right| \\
&= \frac{1}{\hat{g}(b) \cdot g(b)} \left| (\hat{G}(b) - G(b))g(b) + G(b)(g(b) - \hat{g}(b)) \right| \\
&\leq \frac{1}{\hat{g}(b) \cdot g(b)} \left( \left| \hat{G}(b) - G(b) \right| g(b) + G(b) \left| g(b) - \hat{g}(b) \right| \right)
\end{aligned}
$$

Since, $g(b)$ is bounded away from 0 in $C_b(2\epsilon)$, and since $\hat{g}(b)$ converges to $g(b)$, for $m$ larger than some constant, we have that $\hat{g}$ is also bounded away from 0. Thus the multiplier of the above quantity is bounded above by some constant. Then the quantity in the parenthesis is at most $O\left( |\hat{G}(b) - G(b)| + |\hat{g}(b) - g(b)| \right)$. By Lemmas 3 and 4, we then get the result. $\qquad \square$

**Value Trimming.** Since, $C_b(2\epsilon)$ is the map of an interior set $C_v(\zeta)$, for some constant $\zeta$ (with the property that $\zeta \to 0$ as $\epsilon \to 0$), we get that for any fixed interior $C_v(\zeta)$, if $m$ is larger than some constant, then with probability $1 - \delta$:

$$
\sup_{v_t \in C_v(\zeta)} |\hat{v}_t - v_t| \leq O\left( \left( \frac{\log(1/\delta)}{m} \right)^{1/4} \right) \tag{18}
$$

However, there is a small problem, if we apply the inversion formula on all the bids, then we are also applying it to bids that do not fall in the interior $C_b(2\epsilon)$. Since, we have no guarantees about the behavior of $\hat{\xi}$ in that region, it might be that the inversion formula for these bids, will create inverted values $\hat{v}_t$ that are falsely inside the interior $C_v(\zeta)$. Thus we will then falsely count those values in our second stage PDF and CDF estimation. To avoid such spurious insertions of values from such pathological bids that are close to the upper and lower bound of the distribution, we will simply ignore these bids and throw them away.

For the lower end, we simply throw away bids that are smaller than some constant $2\epsilon$. For the upper bound region, there is the technicality, that we don't know the upper bound $\bar{b}$ of the distribution $G$. We will approximate $\bar{b}$ with the maximum observed bid in the data. Observe, that since $g(\cdot)$ is bounded away from zero, for any small interval $[\bar{v} - \kappa, \bar{v}]$, the probability that a bid falls in that interval is bounded away from 0 by some constant $p$. Thus the probability that none of the $m$ bids fall in that region is at most $(1-p)^m \leq e^{-pm}$. For $m \geq \frac{\log(1/\delta)}{p}$, the latter bad event probability is at most $\delta$. Thus with high probability we will have approximated $\bar{b}$ to within an accuracy $\kappa$.

Hence, if we use the inversion formula:

$$
\hat{v}_t = \begin{cases} \hat{\xi}(b_t) & \text{if } b_t \in \left[ 2\epsilon, \max_{\tau \in [m]}\{b_\tau\} - 2\epsilon \right] \\ \infty & \text{o.w.} \end{cases} \tag{19}
$$

Then we have the guarantee, that for some small $\zeta$ (which goes to 0 as $\epsilon$ and $\kappa$ go to zero), all bounded inferred values $\hat{v}_t$ are close to their true values and bounded inferred values correspond to all true values in $C_v(\zeta)$.

## 3.3 Value CDF and PDF Estimation from Inverted Values

Finally, we are ready to show that the small error in the value inversion, cannot lead to a too large of an error in our estimate of the CDF and PDF based on these inverted values $\hat{v}_t$.

We show this for the case of the PDF estimation, $\hat{f}$. We will bound the error by showing that the estimated PDF is close to the un-implementable ideal PDF we would have created if we had the true

values $v_t$. Then the result will follow by triangle inequality and the fact that the ideal PDF is uniformly close to the true PDF, since $f(\cdot)$ is Lipschitz, by assumption.

More formally, let:

$$\tilde{f}_h(v) = \frac{1}{m} \sum_{t=1}^{m} \frac{1}{h} K_u \left( \frac{v - v_t}{h} \right) \tag{20}$$

be the ideal PDF Kernel estimator, where we are using the true values $v_t$ rather than the estimated values $\hat{v}_t$, and we use the uniform Kernel with some bandwidth $h$. We start with a simple Lemma, whose proof is identical to the proof of Lemma 4.

**Lemma 6.** *Under Assumption 2, for any interior subset $C_v(\zeta)$ of the support and for any $h \leq \zeta$ with probability $1 - \delta$:*

$$\sup_{v \in C_v(\zeta)} |\tilde{f}_h(v) - f(v)| \leq \lambda_f h + \frac{1}{h} \frac{O(\log(1/\delta))}{\sqrt{m}} \tag{21}$$

Now we are ready to prove our main result.

**Lemma 7.** *For any fixed interior set $C_v(2\zeta)$, where $\zeta$ is as defined in the previous section, and if we use the uniform Kernel with $h_f \leq \zeta$ and $h_f = O(m^{-1/8})$ and $h_g = O(m^{-1/4})$, then for $m$ larger than some constant, with probability at least $1 - \delta$:*

$$\sup_{v \in C_v(2\zeta)} |\hat{f}(v) - f(v)| \leq O \left( \left( \frac{\log(1/\delta)}{m} \right)^{1/8} \right) \tag{22}$$

**Proof.**    Consider, any value $v \in C_v(2\zeta)$. We compare $\hat{f}(v)$ with $\tilde{f}_h(v)$ for some $h$ and then apply Lemma 6. We deal with the upper and lower bound on the error separately.

**Upper bound on error.**    By trimming, we know that all the finite values that we estimated are within a $\Delta = O \left( \left( \frac{\log(1/\delta)}{m} \right)^{1/4} \right)$ error from their true values. Thus with probability $1 - \delta$ for any $v \in C_v(2\zeta)$:

$$\hat{f}(v) = \frac{1}{m} \sum_{t=1}^{m} \frac{1}{h_f} \mathbf{1}\{|v - \hat{v}_t| \leq h_f\} \leq \frac{1}{m} \sum_{t=1}^{m} \frac{1}{h_f} \mathbf{1}\{|v - v_t| \leq h_f + \Delta\}$$

Since $\hat{v}_t$ is in a region of $h_f$ around $v$ only if $v_t$ is in a region $h_f + \Delta$ around $v$. By multiplying and dividing by $h_f + \Delta$, and applying Lemma 6 we have, with probability $1 - \delta$ for any $v \in C_v(2\zeta)$:

$$
\begin{aligned}
\hat{f}(v) &\leq \frac{h_f + \Delta}{h_f} \frac{1}{m} \sum_{t=1}^{m} \frac{1}{h_f + \Delta} \mathbf{1}\{|v - v_t| \leq h_f + \Delta\} \\
&= \frac{h_f + \Delta}{h_f} \tilde{f}_{h_f + \Delta}(v) && \text{(by definition of } \tilde{f}_h(v)) \\
&\leq \frac{h_f + \Delta}{h_f} \left( f(v) + \lambda_f(h_f + \Delta) + \frac{1}{h_f + \Delta} \frac{O(\log(1/\delta))}{\sqrt{m}} \right) && \text{(by Lemma 6)} \\
&\leq \left( 1 + \frac{\Delta}{h_f} \right) \cdot \left( f(v) + \lambda_f(h_f + \Delta) + \frac{1}{h_f + \Delta} \frac{O(\log(1/\delta))}{\sqrt{m}} \right)
\end{aligned}
$$

By picking $h_f = m^{-1/8}$, we have, that $\frac{\Delta}{h_f} = O \left( \left( \frac{\log(1/\delta}{m} \right)^{1/8} \right)$ and that:

$$\lambda_f(h_f + \Delta) + \frac{1}{h_f + \Delta} \frac{O(\log(1/\delta))}{\sqrt{m}} \leq O \left( \left( \frac{\log(1/\delta}{m} \right)^{1/8} \right) \tag{23}$$

Since $f(v)$ is also upper bounded by a constant, we get that:

$$\hat{f}(v) \leq f(v) + O \left( \left( \frac{\log(1/\delta}{m} \right)^{1/8} \right) \tag{24}$$

**Lower bound on error.** We will use the same setting of $h_f = m^{-1/8}$ and we will show that the density at any point $v \in C_v(2\zeta)$ is lower bounded by the true density, less an error rate of $O(m^{-1/8})$. By definition of $\zeta$, we know that all values $v_t \in C(\zeta)$, are estimated to within an accuracy of $\Delta$. Moreover, since $\zeta \geq h_f \geq \Delta$, for any $v \in C_v(2\zeta)$, and for any $v_t$, such that $|v - v_t| \leq h_f - \Delta$, it must also be that $|v - \hat{v}_t| \leq h_f$. The reason is that such a $v_t$ must fall in $C_v(\zeta)$ and therefore, it must be approximated to within a $\Delta$ accuracy with probability $1 - \delta$.

Thus with probability $1 - \delta$ for any $v \in C_v(2\zeta)$:

$$
\begin{aligned}
\hat{f}(v) =\ & \frac{1}{m}\sum_{t=1}^{m}\frac{1}{h_f}1\{|v - \hat{v}_t| \leq h_f\} \geq\ \frac{1}{m}\sum_{t=1}^{m}\frac{1}{h_f}1\{|v - v_t| \leq h_f - \Delta\} \\
=\ & \frac{h_f - \Delta}{h_f}\frac{1}{m}\sum_{t=1}^{m}\frac{1}{h_f - \Delta}1\{|v - v_t| \leq h_f - \Delta\} \\
=\ & \frac{h_f - \Delta}{h_f}\tilde{f}_{h_f - \Delta}(v) && \text{(by definition of } \tilde{f}_h(v)) \\
\geq\ & \frac{h_f - \Delta}{h_f}\left(f(v) - \lambda_f(h_f - \Delta) - \frac{1}{h_f - \Delta}\frac{O(\log(1/\delta))}{\sqrt{m}}\right) && \text{(by Lemma 6)} \\
\geq\ & \left(1 - \frac{\Delta}{h_f}\right)\cdot\left(f(v) - \lambda_f(h_f - \Delta) - \frac{1}{h_f - \Delta}\frac{O(\log(1/\delta))}{\sqrt{m}}\right)
\end{aligned}
$$

For $h_f = O(m^{-1/8}$, we get that:

$$
\hat{f}(v) \geq f(v) - O\left(\left(\frac{\log(1/\delta)}{m}\right)^{1/8}\right) \tag{25}
$$

$\square$

# 4  Historical Remarks

The result presented in this note, is based on the seminal work of [5] who proposed the two-stage estimation approach for non-parametric inference in first price auctions. A plethora of work has built on top of this work, generalizing to many settings beyond independent valuations and first price auctions. We refer to the recent survey book of [6].

More generally estimation in games has been a vibrant line of work in the past two decades. We refer the reader to the following recent surveys: [3, 1, 2]. You can also find many references and an overview of the literature in the following three-part presentation of a mini-course in Econometric Theory for Games [7, 8, 9].

# References

[1] Daniel Ackerberg, C. Lanier Benkard, Steven Berry, and Ariel Pakes. Econometric Tools for Analyzing Market Outcomes. In J.J. Heckman and E.E. Leamer, editors, *Handbook of Econometrics*, volume 6 of *Handbook of Econometrics*, chapter 63. Elsevier, January 2007.

[2] Patrick Bajari, Han Hong, and Denis Nekipelov. *Game Theory and Econometrics: A Survey of Some Recent Research*, volume 3 of *Econometric Society Monographs*, page 352. Cambridge University Press, 2013.

[3] Steven Berry and Elie Tamer. *Identification in Models of Oligopoly Entry*, volume 2 of *Econometric Society Monographs*, page 4685. Cambridge University Press, 2006.

[4] Shuchi Chawla and Jason D. Hartline. Auctions with unique equilibria. In *Proceedings of the Fourteenth ACM Conference on Electronic Commerce*, EC '13, pages 181–196, New York, NY, USA, 2013. ACM.

[5] Emmanuel Guerre, Isabelle Perrigne, and Quang Vuong. Optimal nonparametric estimation of first-price auctions. *Econometrica*, 68(3):525–574, 2000.

[6] H. J. Paarsch and H. Hong. *An Introduction to the Structural Econometrics of Auction Data*. MIT Press, 2006.

[7] Vasilis Syrgkanis. Econometric theory for games, Part I: Primer and games of incomplete information. `http://vsyrgkanis.com/Econometrics_in_Games_1.pptx`.

[8] Vasilis Syrgkanis. Econometric theory for games, Part II: Dynamic games of incomplete information and games of complete information. `http://vsyrgkanis.com/Econometrics_in_Games_2.pptx`.

[9] Vasilis Syrgkanis. Econometric theory for games, Part III: Auction games and econometrics in algorithmic game theory. `http://vsyrgkanis.com/Econometrics_in_Games_3.pptx`.