Web Supplement to "Inference of Bidders' Risk Attitudes in Ascending Auctions with Endogenous Entry"

Hanming Fang* Xun Tang†

January 20, 2014

1 Probability Mass at $\overline{v}$

This section shows the theoretical, identification and inference results in Fang and Tang (2014) hold when the value distribution is continuous over $[\underline{v}, \overline{v})$ with a probability mass at $\overline{v}$. The model assumptions considered in this section are almost identical to that in Section 3 in Fang and Tang (2014), except for the probability mass at $\overline{v}$.

Assumption M1. *Conditional on K, private values V_i are independent draws from the same continuous marginal distribution $F_{V|K}$ which has positive density almost everywhere with respect to the Lebesgue measure over $[\underline{v}, \overline{v})$ with a probability mass at $\overline{v}$. Entry costs across auctions are independent draws from a continuous distribution F_K with a support $[\underline{k}, \overline{k}]$.*

Let the notations be defined as in the text. To begin with, recall the auction theory suggests that a bidder's dominant strategy in a English auction is still to exit at the true value V_i even in the presence of probability mass at $\overline{v}$ and regardless of the assignment rules when more than one bidders' private values tie at $\overline{v}$.

1.1 Identification

For any generic integrable function g, we use $\int_{\cdot}^{\overline{v}} g(s)ds$ and $\int_{r}^{\cdot} g(s)ds$ as shorthands for the improper integrals $\lim_{v\uparrow\overline{v}} \int_{\cdot}^{v} g(s)ds$ and $\lim_{v\downarrow r} \int_{v}^{\cdot} g(s)ds$ respectively. Let $\omega(k; \boldsymbol{\lambda}_{-i})$ denote expected utility for bidder i conditional on paying entry cost k and potential competitors entering with probabilities $\boldsymbol{\lambda}_{-i} = (\lambda_j)_{j\in N\setminus\{i\}} \equiv (\lambda_1, ., \lambda_{i-1}, \lambda_{i+1}, ., \lambda_N)$. Under Assumption M1,

$$\omega(k; \boldsymbol{\lambda}_{-i}) \equiv u(-k)F_{V|k}(r) + \int_{r}^{\overline{v}} h(v, k; \boldsymbol{\lambda}_{-i})dF_{V|k}(v) + h(\overline{v}, k; \boldsymbol{\lambda}_{-i}) \Pr(V = \overline{v} \mid k) \tag{1}$$

*Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104; and the NBER. Email: hanming.fang@econ.upenn.edu

†Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104. Email: xuntang@econ.upenn.edu

where for all $v \in [r, \overline{v})$,

$$h(v,k;\boldsymbol{\lambda}_{-i}) \equiv u(v-r-k)F_{P_i}(r \mid k,\boldsymbol{\lambda}_{-i}) + \int_r^v u(v-p-k)dF_{P_i}(p \mid k,\boldsymbol{\lambda}_{-i}) + u(-k)[1-F_{Pi}(v \mid k,\boldsymbol{\lambda}_{-i})] \quad (2)$$

with $F_{P_i}(\cdot \mid k,\boldsymbol{\lambda}_{-i})$ being the distribution of P_i when $K=k$, and i's potential competitors enter with probabilities $\boldsymbol{\lambda}_{-i}$. For the boundary case with $v=\overline{v}$, the definition in (2) is adjusted to

$$h(\overline{v},k;\boldsymbol{\lambda}_{-i}) \equiv u(\overline{v}-r-k)F_{P_i}(r \mid k,\boldsymbol{\lambda}_{-i}) + \int_r^{\overline{v}} u(\overline{v}-p-k)dF_{P_i}(p \mid k,\boldsymbol{\lambda}_{-i}) + u(-k)\Pr(P_i=\overline{v} \mid k,\boldsymbol{\lambda}_{-i}) \quad (3)$$

where $\Pr(P_i=\overline{v} \mid k,\boldsymbol{\lambda}_{-i}) = 1-\lim_{\tilde{v}\to\overline{v}} F_{P_i}(\tilde{v} \mid k,\boldsymbol{\lambda}_{-i})$.

It is **important** to note that the form of $h(\overline{v},k;\boldsymbol{\lambda}_{-i})$ in (3) is invariant to the assignment rules used when more than one bidders tie at the upper end of the support $\overline{v}$. This is because when $V_i=\overline{v}$ and $P_i=\overline{v}$, the ex post utility for bidder i is $u(-k)$ regardless of whether he is assigned the object or not.

Due to symmetry in the private value distribution between bidders, $F_{P_i}(.|k,\boldsymbol{\lambda}_{-i})$ does not change with the bidder identity i, and therefore ω does not depend on the bidder identity i. The following lemma characterizes the symmetric mixed strategy equilibrium in the entry stage:

Lemma M1. *Suppose Assumption M1 holds. For any entry cost k with $\omega(k;(1,.,1)) < u(0) < \omega(k;(0,.,0))$, there exists a unique symmetric mixed strategy equilibrium in which all bidders enter with probability λ_k^*, where λ_k^* solves $\omega(k;(\lambda_k^*,.,\lambda_k^*)) = u(0)$.*

Proof of Lemma M1. First, we show that under Assumption M1, $\omega(k;\lambda_{-i})$ is continuous and decreasing in λ_{-i} for all k. To see this, note for any $t \in [r, \overline{v}]$, the event "$P_i \le t$" can be represented as

$$\cap_{j\in N\setminus\{i\}}\{\text{"}j \textit{ stays out}\text{" or "}j \textit{ enters} \cap V_j \le t\text{"}\}.$$

Due to the independence between entry decisions and between private values,

$$F_{P_i}(t \mid k,\boldsymbol{\lambda}_{-i}) = \prod_{j\neq i}[1-\lambda_j+\lambda_j F_{V|k}(t)]. \quad (4)$$

for all $t \in [r,\overline{v}]$. Besides, the probability mass in the distribution of P_i at $v=\overline{v}$ is

$$\Pr(P_i=\overline{v} \mid k,\boldsymbol{\lambda}_{-i}) = 1-\lim_{\tilde{v}\to\overline{v}} F_{P_i}(\tilde{v} \mid k,\boldsymbol{\lambda}_{-i}).$$

The marginal effect of λ_j on this conditional probability in (4) is strictly negative for all $j \neq i$ at $\lambda_j \in [0,1]$ and $t \in [r,\overline{v})$. (Recall r is a binding reserve price.) Thus for all $v \in [r,\overline{v}]$ (including the boundary case with $v=\overline{v}$):

$$h(v,k;\boldsymbol{\lambda}_{-i}) = E[u(v-\tilde{P}_i-k) \mid k,\boldsymbol{\lambda}_{-i}]$$

where the expectation is taken with respect to a random variable $\tilde{P}_i$ whose support is $[r,\overline{v}]$ and whose distribution satisfies $\Pr(\tilde{P}_i=r \mid k,\boldsymbol{\lambda}_{-i}) = F_{P_i}(r \mid k,\boldsymbol{\lambda}_{-i})$ and $\Pr(\tilde{P}_i \le t \mid k,\boldsymbol{\lambda}_{-i}) = F_{P_i}(t \mid k,\boldsymbol{\lambda}_{-i})$ for all $t \in (r,\overline{v})$ and $\Pr(\tilde{P}_i=\overline{v} \mid k,\boldsymbol{\lambda}_{-i}) = \Pr(P_i=\overline{v} \mid k,\boldsymbol{\lambda}_{-i})$. It then follows that $h(v,k;\boldsymbol{\lambda}_{-i})$ is decreasing in $\boldsymbol{\lambda}_{-i}$ given any k and $v \in [r,\overline{v})$. Because it is shown above that the distribution of $\tilde{P}_i$ (which has a probability mass at the upper end of its support) is stochastically increasing λ_{-i} and because $u(.)$ is increasing, it follows that $\omega(k;\boldsymbol{\lambda}_{-i})$ is decreasing in $\boldsymbol{\lambda}_{-i}$. The rest of the proof follows from the same arguments as those used for Lemma 1 in Section 3 of Fang and Tang (2014). □

Whenever $\omega(k;(0,.,0)) \leq u(0)$ (respectively, $\omega(k;(1,.,1)) \geq u(0)$), the equilibrium entry probabilities are degenerate at 0 (respectively, 1). Thus the condition that $\omega(k;(1,.,1)) < u(0) < \omega(k;(0,.,0))$ can be tested, as long as entry decisions are reported in data.

Proposition M1. *Suppose Assumption M1 holds. If K is independent from $(V_i)_{i\in N}$ and $E(K)$ is known to the researcher, then $E[\pi(K)]$ is identified from entry decisions and the distribution of transaction prices.*

Proof of Proposition M1. First off, we show $E[(V_i - P_i)_+ \mid A_{-i} = a] = \int_r^{\bar{v}} F_V(v)^a[1 - F_V(v)]dv$ under Assumption M1 and the other conditions in Proposition M1. By definition, $E[(V_i - P_i)_+ \mid A_{-i} = a]$ is

$$\begin{aligned}
& E[1\{P_i < \bar{v}\}(V_i - P_i)_+ \mid A_{-i} = a] + E[1\{P_i = \bar{v}\} \cdot 0 \mid A_{-i} = a] \\
= \;& E[(V_i - P_i)1\{V_i > P_i > r\}1\{P_i < \bar{v}\} \mid A_{-i} = a] + E[(V_i - P_i)1\{V_i > P_i = r\} \mid A_{-i} = a] \\
= \;& \int_r^{\bar{v}} \phi^*(p) dF_{P_i|A_{-i}=a}(p) + F_V(r)^a \left[\int_r^{\bar{v}} (v - r)\, dF_V(v) + (\bar{v} - r)q_0\right]
\end{aligned}$$

where $q_0 \equiv 1 - \lim_{v\uparrow\bar{v}} F_V(v)$ and $\phi^*(p) \equiv (\bar{v} - p)q_0 + \int_p^{\bar{v}}(v - p)dF_V(v)$. Thus we can write $E[(V_i - P_i)_+ \mid A_{-i} = a]$ as:

$$\begin{aligned}
& \int_r^{\bar{v}} \left[(\bar{v} - p)q_0 + \int_p^{\bar{v}} (v - p)dF_V(v)\right] dF_{P_i|A_{-i}=a}(p) + F_V(r)^a \left[\int_r^{\bar{v}} (v - r)\, dF_V(v) + (\bar{v} - r)q_0\right] \\
= \;& q_0 \left[\int_r^{\bar{v}} (\bar{v} - p)dF_{P_i|A_{-i}=a}(p)\right] + \int_r^{\bar{v}} \left[\int_p^{\bar{v}} (v - p)dF_V(v)\right] dF_{P_i|A_{-i}=a}(p) \qquad (5) \\
& + F_V(r)^a \left[\int_r^{\bar{v}} (v - r)\, dF_V(v) + (\bar{v} - r)q_0\right].
\end{aligned}$$

Using integration by parts, the first term on the R.H.S. of (5) is

$$q_0 \left[0 - (\bar{v} - r)F_{P_i|A_{-i}=a}(r) + \int_r^{\bar{v}} F_{P_i|A_{-i}=a}(p)dp\right]. \qquad (6)$$

By an application of the Bounded Convergence Theorem and integration by parts, the second term on the R.H.S. of (5) is:

$$\begin{aligned}
& \lim_{s\to r^+} \lim_{v'\to\bar{v}^-} \lim_{v''\to\bar{v}^-} \int_s^{v'} \left[\int_p^{v''} (v - p)dF_V(v)\right] dF_{P_i|A_{-i}=a}(p) \\
= \;& \lim_{v'\to\bar{v}^-} \lim_{v''\to\bar{v}^-} \left[\int_{v'}^{v''} (v - v')dF_V(v)\right] F_{P_i|A_{-i}=a}(v') - \lim_{s\to r^+} \lim_{v''\to\bar{v}^-} \left[\int_s^{v''} (v - s)dF_V(v)\right] F_{P_i|A_{-i}=a}(s) \\
& + \lim_{s\to r^+} \lim_{v'\to\bar{v}^-} \lim_{v''\to\bar{v}^-} \int_s^{v'} F_{P_i|A_{-i}=a}(p)[F_V(v'') - F_V(p)]dp \qquad (7)
\end{aligned}$$

where the derivation of the third term follows from the Leibniz Rule and

$$\frac{d}{dp}\left(\int_p^{v''} (v - p)dF_V(v)\right) = 0 - 0 + \int_p^{v''} (-1)dF_V(v) = -[F_V(v'') - F_V(p)].$$

The first term on the R.H.S. of (7) is 0 while the second equals $\left[\int_r^{\bar{v}}(v - r)dF_V(v)\right] F_{P_i|A_{-i}=a}(r)$. By another application of the Bounded Convergence Theorem, the third term on the R.H.S. of (7) is:

$$\lim_{s\to r^+} \lim_{v'\to\bar{v}^-} \int_s^{v'} [F_V(p)]^a\, [\lim_{v''\to\bar{v}} F_V(v'') - F_V(p)]dp = \int_r^{\bar{v}} [F_V(p)]^a\, [1 - q_0 - F_V(p)]dp. \qquad (8)$$

Substituting (6)-(8) into (5) and using $F_{P_i|A_{-i}=a}(p) = F_V(p)^a$ for $p \geq r$, we get:

$$E[(V_i - P_i)_+ \mid A_{-i} = a] = \int_r^{\overline{v}} F_V(v)^a[1 - F_V(v)]dv. \tag{9}$$

The rest of the proof follows from the same steps as in the text. $\square$

1.2 Asymptotic Property of $\hat{\tau}_T$

We sketch the proof of the limiting distribution of the test statistic in almost the same environment as in Section 3 of Fang and Tang (2014), except that there is a probability mass at $\overline{v}$ (as stated in Assumption M1). We also need additional assumptions of independence.

Assumption M2. *(i) The elements in* $(V_1, V_2, ..., V_N, K, \epsilon)$ *are mutually independent with finite second moments and* $E(\epsilon) = 0$*. (ii) The marginal density for* V_i *is the same for all* i *and is bounded above and away from zero by a positive constant. (iii)* $0 < \lambda_k^* < 1$ *for all* $k \in [\underline{k}, \overline{k}]$*.*

Note the tail condition in part (ii) of Assumption 2 in Fang and Tang (2014) is no longer necessary when the distribution of private values assigns a probability mass at $\overline{v}$. Let $\mathbb{P}_0$ denote the true probability measure for (W, A) in data-generating process. Let $\mathbb{P}_T$ denote the corresponding empirical measure. For any signed measure $\mathbb{Q}$, let $\mathbb{Q}f \equiv \int f \, d\mathbb{Q}$ (e.g. $\mathbb{P}_0 1\{W \leq s\} = \Pr(W \leq s)$). Let $F_{W|A=m}$ denote the distribution of W given $A = m$, and $F_{\tilde{W}|A=m}$ denote $\{F_{W|A=m}(t) : t \in [r, \overline{v})\}$. Let $\hat{F}_{\tilde{W}|A=m,T}$ denote estimators for $F_{\tilde{W}|A=m}$ as defined in the text. Let $\hat{\pi}_{T,m} \equiv \mathbb{P}_T 1\{A = m\}$, $\hat{\pi}_T \equiv (\hat{\pi}_{T,m})_{m=0}^N$ and $\pi \equiv (\pi_m)_{m=0}^N$ where $\pi_m \equiv \mathbb{P}_0 1\{A = m\}$. By definition, $F_{W|m}(s) \equiv \Pr(W \leq s \mid A = m) = \pi_m^{-1}\mathbb{P}_0 1\{W \leq s,\, A = m\}$. Recall $\hat{\rho}_T \equiv (\hat{\rho}_{T,a})_{a=0}^{N-1}$ where $\hat{\rho}_{T,a} \equiv \frac{1}{T}\sum_{t \leq T}[\frac{N-a}{N}1\{A_t = a\} + \frac{a+1}{N}1\{A_t = a+1\}]$, and $\rho \equiv (\rho_a)_{a=0}^{N-1}$ where $\rho_a \equiv \Pr(A_{-i} = a)$. The distribution $F_{W,A}$ in the data-generating process (DGP) takes the form of a mixture distribution: $F_{W,A} = \int F_{W,A|k} dF_K$.

The *first* step is to characterize the joint asymptotic property of three building blocks of the test statistic: $\hat{F}_{\tilde{W}|A=m,T}$ for $2 \leq m \leq N$, $\hat{\rho}_T$ and $\frac{1}{T}\sum_{t \leq T} \tilde{K}_t$. For $2 \leq m \leq N$, let $\{\mathbb{F}_{\tilde{W}|A=m} : 2 \leq m \leq N\}$ denote $N-1$ zero-mean Gaussian Processes, each of which is indexed by $[r, \overline{v}]$. The $N-1$ processes are independent with the covariance between $\mathbb{F}_{\tilde{W}|A=m}(s)$ and $\mathbb{F}_{\tilde{W}|A=m'}(s')$ being zero for all $m \neq m'$ and s and $s' \geq r$. Let $(\mathcal{N}_\rho', \mathcal{N}_\mu)'$ denote a multivariate normal random vector in $\mathbb{R}^{N+1}$. For each m, let the covariance between $(\mathcal{N}_\rho', \mathcal{N}_\mu)'$ and $\mathbb{F}_{\tilde{W}|A=m}$ be as specified in Appendix B of Fang and Tang (2014).

Under Assumptions M1 and M2,

$$\sqrt{T}\begin{pmatrix} \hat{F}_{\tilde{W}|A=2,T} - F_{\tilde{W}|A=2} \\ \vdots \\ \hat{F}_{\tilde{W}|A=N,T} - F_{\tilde{W}|A=N} \\ \hat{\rho}_T - \rho \\ \frac{1}{T}\sum_{t \leq T} \tilde{K}_t - \mu_K \end{pmatrix} \rightsquigarrow \begin{pmatrix} \mathbb{F}_{\tilde{W}|A=2} \\ \vdots \\ \mathbb{F}_{\tilde{W}|A=N} \\ \mathcal{N}_\rho \\ \mathcal{N}_\mu \end{pmatrix}.$$

The proof uses the Donsker Property of classes of indicator functions. Define three classes of functions with domain over the support of (W, A):

$$\begin{aligned} \mathcal{F}_W &\equiv \{1\{W \leq s\}\colon s \in [r, \overline{v}]\}; \\ \mathcal{F}_A &\equiv \{1\{A = m\} :\ 0 \leq m \leq N\}; \text{ and} \\ \mathcal{F}_\rho &\equiv \{\tfrac{N-a}{N}1\{A_t = a\} + \tfrac{a+1}{N}1\{A_t = a+1\}\colon 0 \leq a \leq N-1\}. \end{aligned}$$

Let $\mathcal{F}_{W,A} \equiv \{f_W \wedge f_A : f_W \in \mathcal{F}_W, f_A \in \mathcal{F}_A\} \equiv \{1\{W \leq s,\ A = m\}\text{: } s \in [r, \bar{v}],\ 0 \leq m \leq N\}$ denote a class formed by taking the pair-wise infimum of $\mathcal{F}_W$ and $\mathcal{F}_A$. Note $\mathcal{F}_W$ and $\mathcal{F}_A$ are both Donsker Classes. By Theorem 2.10.6 in van der Vaart and Wellner (1996), both $\mathcal{F}_\rho$ and $\mathcal{F}_{W,A}$ are Donsker classes, and consequently $\mathcal{F} \equiv \mathcal{F}_{W,A} \cup \mathcal{F}_\rho \cup \mathcal{F}_A$ is Donsker. For a set $\mathcal{S}$, let $\mathcal{B}(\mathcal{S})$ denote the space of bounded, real-valued functions with domain $\mathcal{S}$, equipped with the sup-norm. By Theorem 2.1 (and the semi-metric defined on page 16) in Kosorok (2008), the empirical process $\mathbb{G}_T \equiv \sqrt{T}(\mathbb{P}_T - \mathbb{P}_0)$ indexed by $\mathcal{F}$ converges weakly to a tight zero-mean Gaussian Process $\mathbb{G}$ in $\mathcal{B}(\mathcal{F})$, with covariance $\mathbb{P}_0 f f' - \mathbb{P}_0 f \mathbb{P}_0 f'$ for any f and $f' \in \mathcal{F}$.

The *second* step is to characterize the joint limiting behavior of $\hat{F}_{V,T}$, $\hat{\rho}_T$ and $T^{-1}\sum_t \tilde{K}_t$. Let $\hat{F}_{V,m,T}(s) \equiv \phi_m^{-1}(\hat{F}_{W|A=m,T}(s))$. To simplify notations, let $F_{\tilde{V}}$ denote $\{F_V(t) : t \in [r, \bar{v})\}$, and similarly let $\hat{F}_{\tilde{V},m,T}$ and $\hat{F}_{\tilde{V},T}$ denote the section of $\hat{F}_{V,m,T}$ and $\hat{F}_{V,T}$ over $[r, \bar{v})$. For each m, let $\xi_m(t) \equiv \phi_m^{-1}(t)$ for $t \in [0, 1)$. For any $s, v \in [r, \bar{v})$ and $m \geq 2$, let $\mathbf{D}_{0,m}$ denote a 2-by-2 diagonal matrix with diagonal entries $\xi_m'\left(F_{W|A=m}(s)\right)$ and $\xi_m'\left(F_{W|A=m}(v)\right)$. Let $\mathbf{\Sigma}_{s,v}$ be a $(2N-2)$-by-$(2N-2)$ block-diagonal matrix such that the $(m-1)$-th diagonal block is the 2-by-2 matrix $\mathbf{D}_{0,m}\tilde{\mathbf{\Sigma}}_{s,v,m}\mathbf{D}_{0,m}'$, where

$$\tilde{\mathbf{\Sigma}}_{s,v,m} \equiv [\Pr(A = m)]^{-1} \begin{pmatrix} F_{W|A=m}(s)\left(1 - F_{W|A=m}(s)\right) & F_{W|A=m}(s)\left(1 - F_{W|A=m}(v)\right) \\ F_{W|A=m}(s)\left(1 - F_{W|A=m}(v)\right) & F_{W|A=m}(v)\left(1 - F_{W|A=m}(v)\right) \end{pmatrix}$$

which is the covariance between $\mathbb{F}_{W|A=m}(s)$ and $\mathbb{F}_{W|A=m}(v)$. Furthermore, define:

$$\underset{2\text{-by-}(2N-2)}{\mathbf{D}_1} \equiv \begin{pmatrix} \frac{1}{N-1} & 0 & \frac{1}{N-1} & 0 & \cdots & \frac{1}{N-1} & 0 \\ 0 & \frac{1}{N-1} & 0 & \frac{1}{N-1} & \cdots & 0 & \frac{1}{N-1} \end{pmatrix}.$$

For any $s \geq r$, let $\mathbf{\Sigma}_s$ denote the variance of the limiting distribution of the random vector $\mathbb{G}_T(1\{W \leq s, A = 2\}, 1\{A = 2\}, ..., 1\{W \leq s, A = N\}, 1\{A = N\}, \tilde{K})$ in $\mathbb{R}^{2N-1}$.Define a 2-by-$(2N-1)$ matrix $\mathbf{D}_2$ as:

$$\begin{pmatrix} \frac{1}{N-1} & \cdots & \frac{1}{N-1} & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} \mathbf{D}_\xi \underset{N\text{-by-}(2N-1)}{\begin{pmatrix} \frac{1}{\pi_2} & \frac{-p_{s2}}{\pi_2\pi_2} & \cdots & 0 & 0 & 0 \\ & & \ddots & & & \\ 0 & 0 & \cdots & \frac{1}{\pi_N} & \frac{-p_{sN}}{\pi_N\pi_N} & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 \end{pmatrix}}$$

where $\mathbf{D}_\xi$ is a N-by-N diagonal matrix with diagonal matrix being $(\xi_2'(F_{W|A=2}(s)), ., \xi_N'(F_{W|A=N}(s)), 1)$ and $p_{sm} \equiv \mathbb{P}_0 1\{W \leq s, A = m\}$ for $s \geq r$.

Under Assumptions M1 and M2,

$$\sqrt{T} \begin{pmatrix} \hat{F}_{\tilde{V},T} - F_{\tilde{V}} \\ \hat{\rho}_T - \rho \\ \frac{1}{T}\sum_t \tilde{K}_t - \mu_K \end{pmatrix} \rightsquigarrow \begin{pmatrix} \mathbb{G}_V \\ \mathcal{N}_\rho \\ \mathcal{N}_\mu \end{pmatrix}$$

where $\mathbb{G}_V$ is a zero-mean Gaussian Process indexed by $[r, \bar{v})$. The covariance kernel for $\mathbb{G}_V$ is $(\mathbb{G}_V(s), \mathbb{G}_V(v)) = \mathbf{D}_1\mathbf{\Sigma}_{s,v}\mathbf{D}_1'$ for any $s, v \in [r, \bar{v})$; and $(\mathbb{G}_V(s), \mathcal{N}_\mu)$ is a bivariate normal with the covariance being $\mathbf{D}_2\mathbf{\Sigma}_s\mathbf{D}_2'$ for any $s \in [r, \bar{v})$. The covariance between $\mathbb{G}_V$ and $\mathcal{N}_\rho$ can also be characterized in a similar fashion using the Delta method. Proof of these results follow from the differentiability of ξ over $[r, \bar{v})$ and an application of the Multi-variate Delta Method, and is omitted. Note that the delta method can be applied here because $\xi'(.)$ is bounded above over $[r, \bar{v})$ due to the probability mass at $\bar{v}$.

The *final* step is to apply the Functional Delta Method to derive the limiting distribution of $\sqrt{T}(\hat{\tau}_T - \tau_0)$. Recall $\mathcal{B}_{[r,\bar{v})}$ denotes the set of positive, bounded and integrable Cadlag functions over $[r, \bar{v})$.

Equipped with a sup-norm, $\mathcal{B}_{[r,\bar{v})}$ is normed linear spaces with a non-degenerate interior. It follows from Lemma 20.10 in van der Vaart (1998) that the mapping ζ as defined in the text is Hadamard differentiable at $F_{\tilde{V}}$ tangentially to $\mathcal{B}_{[r,\bar{v})}$, with the derivative $D_{\zeta,F_{\tilde{V}}} : \mathcal{B}_{[r,\bar{v})} \to \mathbb{R}_+^N$ being

$$D_{\zeta,F_{\tilde{V}}}(h)(a) \equiv \int_r^{\bar{v}} \left\{ a[F_{\tilde{V}}(s)]^{a-1} h(s) - (a+1)[F_{\tilde{V}}(s)]^a h(s) \right\} ds \tag{10}$$

for any $h \in \mathcal{B}_{[r,\bar{v})}$ and $0 \leq a \leq N-1$. It is *important* to note that the integral in (9) and (10) are improper under Assumption M1. That is, the integrals $\int^{\bar{v}}$ are defined as $\lim_{\tilde{v}\to\bar{v}} \int^{\tilde{v}}$. Hence the right-hand sides of (9) and (10) are both operators defined over $\mathcal{B}_{[r,\bar{v})}$.

The Jacobian of τ with respect to its components $(\zeta,\ \rho,\ \mu_K)$ at the their true DGP values is:

$$\left(\rho_0, ., \rho_{N-1}, \zeta_0, ., \zeta_{N-1}, -1\right) \equiv (\rho, \zeta, -1).$$

Since ζ is Hadamard differentiable at $F_{\tilde{V}}$ tangentially to $\mathcal{B}_{[r,\bar{v})}$,[1] it follows from the Functional Delta Method (Theorem 2.8 in Kosorok (2008)) that

$$\sqrt{T} \begin{pmatrix} \hat{\zeta}_T - \zeta \\ \hat{\rho}_T - \rho \\ \frac{1}{T}\sum_t \tilde{K}_t - \mu_K \end{pmatrix} \rightsquigarrow \begin{pmatrix} D_{\zeta,F_{\tilde{V}}}(\mathbb{G}_V) \\ \mathcal{N}_\rho \\ \mathcal{N}_\mu \end{pmatrix}.$$

An application of the multivariate delta method shows under Assumptions M1 and M2, $\sqrt{T}\,(\hat{\tau}_T - \tau_0) \rightsquigarrow \mathcal{N}_\tau \equiv \rho' D_{\zeta,F_{\tilde{V}}}(\mathbb{G}_V) + \zeta' \mathcal{N}_\rho - \mathcal{N}_\mu$. To see that the limiting distribution $\mathcal{N}_\tau$ is univariate normal with zero-mean, note the Gaussian process $\mathbb{G}_V$ is Borel-measurable and tight (see Example 1.7.3. in van der Vaart and Wellner (1996)) and that by construction the Hadamard derivative $D_{\zeta,F_{\tilde{V}}}$ is a linear mapping defined over $\mathcal{B}_{[r,\bar{v})}$. It follows from Lemma 3.9.8. of van der Vaart and Wellner (1996) that $\mathcal{N}_\tau$ is univariate normal with zero mean.

1.3 Monte Carlo Simulation

In this section we present Monte Carlo evidence for the performance of our test in finite samples. We consider the following data-generating process (DGP). Each auction involves N potential bidders who face a common entry cost K. Upon entry, bidders draw private values from the support $[\underline{v}, \bar{v}] = [0, 10]$. There is a probability mass of 2% at $\bar{v}$ and the rest of the probability mass is uniformly spread over the half-open interval $[\underline{v}, \bar{v})$. The reserve price is set at $r = 3$. The data contains prices paid by the winners and the number of entrants in each auction. In the rare case where the reserve price screens out all entrants (i.e. privates values for all entrants are lower than r), transaction prices are defined to be an arbitrary number smaller than r. The entry cost in each auction is drawn from a multinomial distribution over the support $\{0.7, 0.8, 0.9\}$ with identical probability masses. The data reports noisy measures of entry costs $\tilde{K} = K + \epsilon$ but not K, where ϵ is drawn from a uniform distribution $[-0.5, 0.5]$. Bidders' von-Neumann-Morgenstern utility is specified as $u(c) \equiv \left(\frac{c+5}{10}\right)^\gamma$, so that they are risk-neutral (or respectively, risk-averse or risk-loving) if $\gamma = 1$ (or $\gamma < 1$ or $\gamma > 1$).[2] We experiment with the number

[1] It is crucial that the domain of the space of functions here is $[r, \bar{v})$ as opposed to $[r, \bar{v}]$.

[2] The specification in this section differs slightly from that used in Section 5 of Fang and Tang (2014). This is mostly due to the fact that the specified supports of private values are different between these two sections. We adopt these utility specifications so that for any given γ the curvature of the utility function over the respective supports of V are comparable across these two sections.

of potential bidders $N = 4$ or $N = 5$, and sample sizes $T = 1,500$ or $T = 3,000$.[3]

To improve the test performance in small samples, we modify our estimator for risk-premium $\hat{\tau}_T$ slightly by replacing the average of estimators for the value distribution F_V in Section 4.2 of Fang and Tang (2014) with inverse-variance-weighted average estimators. That is, while constructing $\hat{\tau}_T$, we replace $\hat{F}_{V,T}(s)$ in Section 4.2 of Fang and Tang (2014) by $\tilde{F}_{V,T}(s) \equiv \frac{1}{N-1}\sum_{m=2}^{N} \hat{\beta}_m(s)\phi_m^{-1}(\hat{F}_{W|A=m,T}(s))$ for $s \geq r$, with $\hat{\beta}_m(s) \equiv \left(\hat{\sigma}^2_{m,T}(s)\right)^{-1} / \sum_{m'=2}^{N}\left(\hat{\sigma}^2_{m',T}(s)\right)^{-1}$ and $\hat{\sigma}^2_{m,T}$ being the standard errors for the estimator of F_V using observations with m entrants. The test statistic then uses the weighted version $\tilde{F}_{V,T}$ in subsequent steps. As the test statistic remains a smooth function of sample averages, we use bootstrap estimates for critical values in its sampling distribution.

[Insert Figure 1 (a), (b), (c), (d) here.]

Panels in Figure 1 report test performance under various DGP and sample sizes T. For each grid point of γ between 0.75 and 1.25 (with a grid-width of 0.05), we calculate an integrated measure of relative risk-aversion $\theta(\gamma)$, defined as integral of $-cu''(c)/u'(c)$ over the support of $c = (V_i - P_i)_+ - K$. (Thus positive values for $\theta(\gamma)$ correspond to risk-averse and negative values to risk-loving bidders.) For each γ and sample size $T = 1,500$ or $3,000$, we simulate $S = 250$ samples.

For each simulated sample, we calculate the statistic $\hat{\tau}_T$, and then perform the test using critical values estimated from $B = 300$ bootstrap samples drawn from this simulated sample. We reports test results for significance levels $\alpha = 5\%$, 10% and 15%. Solid curves in each panel show the proportions of S samples in which our test fails to reject the null of risk-neutrality ($H_0 : \theta(\gamma) = 0$). The dashed curves (and dotted curves respectively) plot the proportions of S samples in which the test rejects the null in favor of alternatives $H_A : \theta(\gamma) > 0$ (and $H_L : \theta(\gamma) < 0$ respectively). Each panel of Figure 1 reports the result for a given pair of T and N, with the integrated measure of risk-aversion $\theta(\gamma)$ plotted on the horizontal axis.

In all panels of Figure 1, the test approximately attains targeted levels under the null. Also, in all panels, the power of our test approaches 1 as the absolute value of the integrated measure of risk-aversion $\theta(\gamma)$ increases.[4] The comparison of panel (a) with (b) and the comparison of panel (c) with (d) suggest, as sample sizes increase, errors in rejection probabilities under the null decrease while the power under any fixed alternatives increases.

Table 1 further quantifies these changes in test performance for $\gamma \in \{0.8, 0.9, 1.0, 1.1, 1.2\}$ and $N \in \{4, 5\}$ by reporting numerical results. Each row corresponds to some pair of (N, γ) and some sample size T, while the column headings are targeted significance levels. For each cell in Table 1 we report, from the left to the right, the proportions of S simulated samples, where the test rejects H_0 in favor of H_L, where the test does not reject H_0, and where H_0 is rejected in favor of H_A respectively.

Table 1(a) shows results for auctions with four potential bidders. Even with a moderate sample size $T = 1,500$, the test attains rejection probabilities that are reasonably close to the targeted level α under the null, and reasonably high probabilities for rejecting the null in favor of the correct alternative when

[3] We use a smaller sample size T compared to those in Section 5 of Fang and Tang (2014). This is because in the latter case we need sufficient observations given each pair (z, n) in the estimation of entry probabilities.

[4] In the presence of directional alternatives (i.e. H_A and H_L), we define the power of the test as the probability of rejecting the null in favor of the true alternative in the DGP.

$\gamma \neq 1$. When $\gamma \neq 1$, the probability for "Type-III" error (i.e. rejecting the null in favor of a wrong alternative) is practically zero across all specifications and sample sizes.

Table 1(a): Probabilities for Accepting $[H_L, H_0, H_A]$: $(N = 4)$

	$\alpha = 5\%$	$\alpha = 10\%$	$\alpha = 15\%$
$T = 1,500$			
$\gamma = 0.8$	[0.004, 0.264, 0.732]	[0.004, 0.192, 0.804]	[0.004, 0.152, 0.844]
$\gamma = 0.9$	[0.008, 0.768, 0.224]	[0.008, 0.676, 0.316]	[0.028, 0.584, 0.388]
$\gamma = 1$	[0.052, 0.932, 0.016]	[0.092, 0.872, 0.036]	[0.132, 0.808, 0.060]
$\gamma = 1.1$	[0.272, 0.724, 0.004]	[0.420, 0.572, 0.008]	[0.512, 0.480, 0.008]
$\gamma = 1.2$	[0.672, 0.328, 0.000]	[0.800, 0.200, 0.000]	[0.880, 0.120, 0.000]
$T = 3,000$			
$\gamma = 0.8$	[0.000, 0.048, 0.952]	[0.000, 0.036, 0.964]	[0.000, 0.020, 0.980]
$\gamma = 0.9$	[0.000, 0.540, 0.460]	[0.004, 0.456, 0.540]	[0.004, 0.388, 0.608]
$\gamma = 1$	[0.044, 0.936, 0.020]	[0.076, 0.888, 0.036]	[0.132, 0.820, 0.048]
$\gamma = 1.1$	[0.548, 0.452, 0.000]	[0.688, 0.312, 0.000]	[0.776, 0.224, 0.000]
$\gamma = 1.2$	[0.972, 0.028, 0.000]	[0.992, 0.008, 0.000]	[1.000, 0.000, 0.000]

Table 1(b): Probabilities for Accepting $[H_L, H_0, H_A]$: $(N = 5)$

	$\alpha = 5\%$	$\alpha = 10\%$	$\alpha = 15\%$
$T = 1,500$			
$\gamma = 0.8$	[0.004, 0.480, 0.516]	[0.008, 0.384, 0.608]	[0.008, 0.324, 0.668]
$\gamma = 0.9$	[0.008, 0.896, 0.096]	[0.016, 0.812, 0.172]	[0.032, 0.732, 0.236]
$\gamma = 1$	[0.064, 0.924, 0.012]	[0.068, 0.916, 0.016]	[0.104, 0.876, 0.020]
$\gamma = 1.1$	[0.192, 0.808, 0.000]	[0.300, 0.700, 0.000]	[0.400, 0.600, 0.000]
$\gamma = 1.2$	[0.620, 0.380, 0.000]	[0.760, 0.240, 0.000]	[0.840, 0.160, 0.000]
$T = 3,000$			
$\gamma = 0.8$	[0.000, 0.096, 0.904]	[0.000, 0.064, 0.936]	[0.000, 0.056, 0.944]
$\gamma = 0.9$	[0.008, 0.588, 0.404]	[0.016, 0.468, 0.516]	[0.016, 0.412, 0.572]
$\gamma = 1$	[0.040, 0.952, 0.008]	[0.076, 0.888, 0.036]	[0.124, 0.828, 0.048]
$\gamma = 1.1$	[0.456, 0.544, 0.000]	[0.588, 0.412, 0.000]	[0.668, 0.332, 0.000]
$\gamma = 1.2$	[0.920, 0.080, 0.000]	[0.992, 0.008, 0.000]	[0.992, 0.008, 0.000]

With the sample size as small as $T = 1,500$ or $3,000$, the power of the tests appears low at $\gamma = 0.9$ or $\gamma = 1.1$. We argue that this should not be interpreted as evidence of unsatisfactory finite sample performance of our test. Rather it is mostly due to the fact that the curvature of utility functions are close to being linear for $\gamma = 0.9$ or $\gamma = 1.1$. The power does improve substantially either as the sample size increases, or as γ moves further away from 1.

Table 1(a) also quantifies the improvement of test performance as the sample size T increases. For a fixed level of α, increasing sample sizes from $T = 1,500$ to $T = 3,000$ reduces the error in the rejection probability by a small amount under the null with $\gamma = 1$. On the other hand, such an increase yields more substantial improvements in the power under each fixed alternatives. Table 1(b) reports results for auctions involving five potential bidders. Overall, it registers the same pattern as in Table 1(a).

Comparisons between Table 1(a) and 1(b) suggest the impact of a larger number of potential bidders on the errors in rejection probabilities under the null is ambiguous. On the other hand, the impact of a large N on the power of the test under any fixed alterative seems unambiguous: For all $\gamma \neq 1$ and given any sample size, the power appears to be higher for all α when $N = 4$. Such a difference in comparison is related to the following fact: Under any alternative, the magnitude of the risk-premium is partly determined by the shape of the distribution of $(V_i - P_i)_+$ given N ; on the other hand this magnitude remains fixed at 0 as the distribution of $(V_i - P_i)_+$ changes with N under the null. Since the size of risk premia, subject to estimation errors in data, essentially determines how likely it is to detect non-risk-neutrality, we conjecture this difference in comparative statics might explain the pattern above.

1.4 Mismeasurement of Entry Costs

This subsection provides further simulation evidence about the test performance when entry costs K are mismeasured. We report results in a DGP where the noisy measures of entry costs systematically understate the true entry costs known to potential bidders. The specification of the DGP is the same as that used above, except that the noisy measures $\tilde{K}$ now has a downward bias equal to $\tilde{\delta}$-percent of the true entry cost K. We experiment with $\tilde{\delta} = 5$ or 10 in simulations. Table 2 reports results with $S = B = 200$ (where S is the number of simulate data sets and B is the number of bootstrap samples used for estimating critical values in each simulated sample).

Table 2(a): Probabilities for Accepting $[H_L, H_0, H_A]$: $(\tilde{\delta} = 5, N = 4)$

	$\alpha = 5\%$	$\alpha = 10\%$	$\alpha = 15\%$
$T = 1,500$			
$\gamma = 0.9$	[0.000, 0.530, 0.470]	[0.000, 0.425, 0.575]	[0.000, 0.380, 0.620]
$\gamma = 1.0$	[0.025, 0.830, 0.145]	[0.040, 0.750, 0.210]	[0.040, 0.680, 0.280]
$\gamma = 1.1$	[0.145, 0.855, 0.000]	[0.225, 0.775, 0.000]	[0.315, 0.685, 0.000]
$\gamma = 1.2$	[0.460, 0.540, 0.000]	[0.600, 0.400, 0.000]	[0.710, 0.290, 0.000]
$T = 3,000$			
$\gamma = 0.9$	[0.000, 0.205, 0.795]	[0.000, 0.125, 0.875]	[0.000, 0.075, 0.925]
$\gamma = 1.0$	[0.015, 0.915, 0.070]	[0.050, 0.820, 0.130]	[0.065, 0.770, 0.165]
$\gamma = 1.1$	[0.180, 0.820, 0.000]	[0.255, 0.745, 0.000]	[0.350, 0.650, 0.000]
$\gamma = 1.2$	[0.850, 0.150, 0.000]	[0.945, 0.055, 0.000]	[0.955, 0.045, 0.000]

Table 2(b): Probabilities for Accepting $[H_L, H_0, H_A]$: $(\tilde{\delta} = 10, N = 4)$

	$\alpha = 5\%$	$\alpha = 10\%$	$\alpha = 15\%$
$T = 1,500$			
$\gamma = 0.9$	[0.000, 0.285, 0.715]	[0.000, 0.205, 0.795]	[0.000, 0.175, 0.825]
$\gamma = 1.0$	[0.005, 0.615, 0.380]	[0.005, 0.530, 0.465]	[0.005, 0.460, 0.535]
$\gamma = 1.1$	[0.090, 0.905, 0.005]	[0.145, 0.845, 0.010]	[0.175, 0.790, 0.035]
$\gamma = 1.2$	[0.355, 0.645, 0.000]	[0.450, 0.550, 0.000]	[0.525, 0.475, 0.000]
$T = 3,000$			
$\gamma = 0.9$	[0.000, 0.040, 0.960]	[0.000, 0.025, 0.975]	[0.000, 0.020, 0.980]
$\gamma = 1.0$	[0.015, 0.815, 0.170]	[0.020, 0.700, 0.280]	[0.025, 0.605, 0.370]
$\gamma = 1.1$	[0.055, 0.940, 0.005]	[0.085, 0.895, 0.020]	[0.110, 0.860, 0.030]
$\gamma = 1.2$	[0.590, 0.410, 0.000]	[0.735, 0.265, 0.000]	[0.835, 0.165, 0.000]

As expected, the two panels in Table 2 show that, compared with Table 1, underestimating the mean of entry costs results in larger errors in rejection probabilities under the null. For instance, for $\tilde{\delta} = 5$ and $\alpha = 5\%$, the probability for rejecting the null is 17% (of which 14.5% leads to conclusions in favor of the risk-averse alternative H_A) when $T = 1500$; and is 8.5% (of which 7% is in favor of H_A) when $T = 3000$. Also the power against risk-loving alternatives is lower than in Table 1 due to the underestimation of entry costs, while that against the risk-averse alternative becomes higher than in Table 1. Similar patterns persist for tests with higher significance levels $\alpha = 10\%$ and $\alpha = 15\%$.

A comparison between the two panels of Table 2 suggests that, as the magnitude of bias increases, the performance of the test worsens both in the sense of larger errors in rejection probabilities and lower power against the risk-loving alternatives. Besides, the probability for Type III error could be positive for tests with higher significance levels and larger bias. With a large bias $\tilde{\delta} = 10$ and $\gamma = 1.1$, the power actually decreases with sample sizes. This can be explained by the fact that the bias in estimating entry costs is large relative to the difference in risk premia under the null and the risk-loving alternative $\gamma = 1.1$.

By construction, the errors in projection probabilities under H_0 due to the underestimation of expected entry costs will not diminish to zero as the sample size approaches infinity. Also, for a range of risk-loving alternatives that are sufficiently close to the null, the test may be inconsistent (the probabilities for rejecting the null given those alternatives do not approach 1 as the sample size increases). However, on a more positive note, for risk-loving alternatives farther away from the null (e.g. $\gamma \geq 1.2$), the test appears to be consistent, with the power against any such alternative approaching 1 as the sample size increases. In sum, the impact on test performance under mismeasurement of entry costs depends on the size of the bias in the estimation of entry costs as well as the distance between the null and the alternatives (as measured by the difference between the risk premia under the null and the alternative).

2 Proof of Lemma B4

As in Appendix B of Fang and Tang (2014), let $\mathcal{B}_{[\underline{v},\overline{v}]}$ denote the space of bounded functions with domain $[\underline{v}, \overline{v}]$; and $\mathcal{D}_{[\underline{v},\overline{v}]} \subset \mathcal{B}_{[\underline{v},\overline{v}]}$ denote the space of bounded, non-decreasing functions that are right-continuous with left limits and map from $[\underline{v}, \overline{v}]$ into $[0, 1]$. Throughout this section, convergence of sequences over $\mathcal{B}_{[\underline{v},\overline{v}]}$ is defined in terms of the sup-norm. For any $1 \leq a \leq N - 1$ and $2 \leq m \leq N$, define $\chi_{a,m}(t) \equiv [\phi_m^{-1}(t)]^a$ for $t \in [0, 1]$ and ϕ_m^{-1} is the mapping from the distribution of a second-largest order statistic from m

independent draws to the parent distribution from which the draws are made. The following lemma summarizes some useful properties of $\chi_{a,m}$.

Lemma W1. *For all* $1 \le a \le N-1$ and $2 \le m \le N$, *(i)* $\chi_{a,m}$ *is continuous and increasing over* $[0,1]$ *with* $\chi_{a,m}(0)=0$ *and* $\chi_{a,m}(1)=1$; *(ii) there exists* $t_{a,m} \in [0,1)$ *such that* $\chi_{a,m}$ *is convex over* $[t_{a,m},1]$; *and (iii)*

$$\lim_{t' \to 1^-} \left(\frac{\chi_{a,m}(1) - \chi_{a,m}(t')}{\sqrt{1-t'}} \right) = a\mathcal{C}_m$$

where $\mathcal{C}_m$ *is a finite, positive constant that depends on* m.

Proof of Lemma W1. By construction, $\phi_m^{-1}(t)$ is strictly increasing over $[0,1]$ with $\phi_m^{-1}(0)=0$ and $\phi_m^{-1}(1)=1$ for all $2 \le m \le N$. This implies (i) for all $a \ge 1$ and $m \ge 2$. Next, note the inverse of $\chi_{a,m}$, denoted $\chi_{a,m}^{-1}$, is

$$\chi_{a,m}^{-1}(s) \equiv \phi_m(s^{\frac{1}{a}}) = (s^{\frac{1}{a}})^m + m(s^{\frac{1}{a}})^{m-1}(1-s^{\frac{1}{a}}) = (1-m)s^{\frac{m}{a}} + ms^{\frac{m-1}{a}}$$

for any $s \in [0,1]$. The second-order derivative of $\chi_{a,m}^{-1}$ at $s \in (0,1)$ is negative if and only if $(m-a)s^{\frac{1}{a}} > m-a-1$. If $m > a+1$, $\chi_{a,m}^{-1}$ is concave over $s > \left(\frac{m-a-1}{m-a}\right)^a$. Or equivalently, $\chi_{a,m}$ is convex over the interval ($\chi_{a,m}^{-1}(\left(\frac{m-a-1}{m-a}\right)^a)$, 1]. If $m \le a+1$, $\chi_{a,m}^{-1}$ is concave over $[0,1]$ (or equivalently $\chi_{a,m}$ is convex over $[0,1]$. Thus (ii) holds for all $a \ge 1$ and $m \ge 2$. As for part (iii), it follows from Menzel and Morganti (2013) that

$$\lim_{t \to 1^-} \frac{\chi_{a,m}(1) - \chi_{a,m}(t)}{\sqrt{1-t}} = \lim_{t \to 1^-} \frac{\chi_{a,m}(1) - \chi_{a,m}(t)}{\phi_m^{-1}(1) - \phi_m^{-1}(t)} \times \lim_{t \to 1^-} \frac{\phi_m^{-1}(1) - \phi_m^{-1}(t)}{\sqrt{1-t}} = a \times \mathcal{C}_m$$

with $\mathcal{C}_m$ being a positive finite constant that depends on m. (See Proposition 3.1 in Menzel and Morganti (2013) with $K \equiv m$, $k \equiv m-1$.) Thus (iii) holds for all $a \ge 1$ and $m \ge 2$. $\square$

Lemma W2. *Suppose Assumptions 1 and 2 in Fang and Tang (2014) hold. Then for all* $2 \le m \le N$, $F_{W|A=m}$ *is continuous and increasing over* $[r, \overline{v}]$ *and has bounded positive density over* $[r, \overline{v})$ *with* $\int_r^{\overline{v}} [1 - F_{W|A=m}(v)]^{-\frac{1}{2}} dv < \infty$.

Proof of Lemma W2. The continuity and monotonicity of $F_{W|A=m}$ holds because $F_{W|A=m} = \phi_m(F_V)$ for all $m \ge 2$ where $\phi_m(.)$ is continuous and increasing over $[0,1]$. That its density is positive and bounded above over $[r, \overline{v})$ follows from the restrictions on the distribution of V in Assumptions 1 and 2 in Fang and Tang (2014) and the fact that $F'_{W|A=m}(v) = m(m-1)F_V(v)^{m-2}F'_V(v)[1-F_V(v)]$ and $F_V(r) > 0$. By construction, $F_{W|A=m}$ is stochastically ordered in m with $F_{W|A=m}(v) \le F_{W|A=2}(v)$ for all $v \in [\underline{v}, \overline{v}]$ and $m \ge 2$.[5] Hence $\int_r^{\overline{v}} [1 - F_{W|A=m}(v)]^{-\frac{1}{2}} dv$ is bounded above by $\int_r^{\overline{v}} [1 - F_{W|A=2}(v)]^{-\frac{1}{2}} dv$ for all $m \ge 2$. By construction, $[1 - F_{W|A=2}(v)]^{-\frac{1}{2}} = [1 - 2F_V(v) + (F_V(v))^2]^{-\frac{1}{2}} = [1 - F_V(v)]^{-1}$. It then follows from (ii) in Assumption 2 in Fang and Tang (2014) that

$$\int_r^{\overline{v}} [1 - F_{W|A=m}(v)]^{-\frac{1}{2}} dv \le \int_r^{\overline{v}} [1 - F_{W|A=2}(v)]^{-\frac{1}{2}} dv = \int_r^{\overline{v}} [1 - F_V(v)]^{-1} dv < \infty.$$

This proves the claim in Lemma W2. $\square$

[5] To see this, note $\phi_{m+1}(t) - \phi_m(t) = -mt^{m-1}(1-t)^2 \le 0$ for all $m \ge 2$ and $t \in [0,1]$.

Let $\chi'_{a,m}(c) \equiv a\left[\phi_m^{-1}(c)\right]^{a-1}\left(d\phi_m^{-1}(t)/dt\big|_{t=c}\right)$; and let $\hat{\mathcal{D}}_{[\underline{v},\overline{v}]} \subset \mathcal{D}_{[\underline{v},\overline{v}]}$ denote the subset of $\mathcal{D}_{[\underline{v},\overline{v}]}$ which consists of step functions only. Assumptions 1 and 2 in Fang and Tang (2014) imply that for any sequence $F_{W|A=m,T} \in \hat{\mathcal{D}}_{[\underline{v},\overline{v}]}$ and for all T, the set $\{v : r \leq v \leq \overline{v}$ and $F_{W|A=m,T}(v) = F_{W|A=m}(v)\}$, which has no contribution to the integral $\int_r^{\overline{v}} \sqrt{T}[\chi_{a,m}(F_{W|A=m,T}(v)) - \chi_{a,m}(F_{W|A=m}(v))]dv$, must have has zero Lebesgue measure.

Lemma W3. *Suppose Assumptions 1 and 2 in Fang and Tang (2014) hold. For any $a \geq 1$, $m \geq 2$ and any $\varepsilon > 0$, there exists $\delta > 0$ so that the following statement holds: "For any sequence $F_{W|A=m,T} \in \hat{\mathcal{D}}_{[\underline{v},\overline{v}]}$ such that $F_{W|A=m,T} \to F_{W|A=m}$ uniformly over $[\underline{v},\overline{v}]$,*

$$\int_{\overline{v}-\delta}^{\overline{v}} \left(\frac{\chi_{a,m}(F_{W|A=m,T}(v)) - \chi_{a,m}(F_{W|A=m}(v))}{F_{W|A=m,T}(v) - F_{W|A=m}(v)}\right) dv < \varepsilon \tag{11}$$

when T is sufficiently large".

Proof of Lemma W3. Lemma W1 implies that, for a positive constant η sufficiently small and for all $v \in [\overline{v}-\eta, \overline{v}]$, both $F_{W|A=m}(v)$ and $F_{W|A=m,T}(v)$ lie in an interval over which $\chi_{a,m}$ is convex, provided T is sufficiently large. In other words, there exists $\delta_1 > 0$ such that $\min\{F_{W|A=m}(v),\, F_{W|A=m,T}(v)\} > t_{a,m}$ for all $v \in [\overline{v}-\delta_1, \overline{v}]$ when T is sufficiently large (where $t_{a,m}$ is defined in part (ii) of Lemma W1). This implies for all $v \in [\overline{v}-\delta_1, \overline{v}]$ such that $F_{W|A=m,T}(v) \neq F_{W|A=m}(v)$,

$$\begin{aligned}
& \frac{\chi_{a,m}(F_{W|A=m,T}(v)) - \chi_{a,m}(F_{W|A=m}(v))}{F_{W|A=m,T}(v) - F_{W|A=m}(v)} \\
\leq \quad & \frac{\chi_{a,m}(1) - \chi_{a,m}(F_{W|A=m}(v))}{1 - F_{W|A=m}(v)} \\
= \quad & \left(\frac{\chi_{a,m}(1) - \chi_{a,m}(F_{W|A=m}(v))}{\sqrt{1 - F_{W|A=m}(v)}}\right)\left(\frac{1}{\sqrt{1 - F_{W|A=m}(v)}}\right)
\end{aligned}$$

when T is sufficiently large. Monotonicity of $\chi_{a,m}$ over $[r, \overline{v}]$ implies $\frac{\chi_{a,m}(F_{W|A=m,T}(v)) - \chi_{a,m}(F_{W|A=m}(v))}{F_{W|A=m,T}(v) - F_{W|A=m}(v)}$ must be positive. Besides, the property in (iii) of Lemma W1 implies for all $\varepsilon' > 0$, there exists $\delta_2 > 0$ such that

$$\max\{0, a\mathcal{C}_m - \varepsilon'\} \leq \frac{\chi_{a,m}(1) - \chi_{a,m}(F_{W|A=m}(v))}{\sqrt{1 - F_{W|A=m}(v)}} \leq a\mathcal{C}_m + \varepsilon'$$

for all $v \in [\overline{v}-\delta_2, \overline{v}]$. For any $\delta \leq \min\{\delta_1, \delta_2\}$,

$$0 \leq \int_{\overline{v}-\delta}^{\overline{v}} \left(\frac{\chi_{a,m}(F_{W|A=m,T}(v)) - \chi_{a,m}(F_{W|A=m}(v))}{F_{W|A=m,T}(v) - F_{W|A=m}(v)}\right) dv \leq (a\mathcal{C}_m + \varepsilon')\left(\int_{\overline{v}-\delta}^{\overline{v}} \left[1 - F_{W|A=m}(v)\right]^{-\frac{1}{2}} dv\right)$$

for T sufficiently large. With $a\mathcal{C}_m + \varepsilon'$ being positive and finite, the claim in this lemma follows from Lemma W2 for δ sufficiently small. $\square$

Lemma W4. *Suppose Assumptions 1 and 2 in Fang and Tang (2014) hold. For any $a \geq 1$, $m \geq 2$ and any $F_{W|A=m,T} \in \hat{\mathcal{D}}_{[\underline{v},\overline{v}]}$ such that $\sqrt{T}(F_{W|A=m,T} - F_{W|A=m}) \to F^*$ for some $F^* \in \mathcal{B}_{[\underline{v},\overline{v}]}$,*

$$\lim_{T\to\infty} \int_r^{\overline{v}} \left(\frac{\chi_{a,m}(F_{W|A=m,T}(v)) - \chi_{a,m}(F_{W|A=m}(v))}{F_{W|A=m,T}(v) - F_{W|A=m}(v)}\right) dv = \int_r^{\overline{v}} \chi'_{a,m}(F_{W|A=m}(v))dv. \tag{12}$$

Proof of Lemma W4. We need to show that for any $\varepsilon > 0$ and any sequence $F_{W|A=m,T} \in \hat{\mathcal{D}}_{[\underline{v},\overline{v}]}$ such that $\sqrt{T}(F_{W|A=m,T} - F_{W|A=m}) \to F^*$ for some $F^* \in \mathcal{B}_{[\underline{v},\overline{v}]}$,

$$\left| \int_r^{\overline{v}} \left(\frac{\chi_{a,m}(F_{W|A=m,T}(v)) - \chi_{a,m}(F_{W|A=m}(v))}{F_{W|A=m,T}(v) - F_{W|A=m}(v)} - \chi'_{a,m}(F_{W|A=m}(v)) \right) dv \right| < \varepsilon$$

when T is sufficiently large.

We first show there exists a constant $\delta(\varepsilon) > 0$ such that for any such sequence $F_{W|A=m,T} \in \hat{\mathcal{D}}_{[\underline{v},\overline{v}]}$,

$$\left| \int_{\overline{v}-\delta(\varepsilon)}^{\overline{v}} \left(\frac{\chi_{a,m}(F_{W|A=m,T}(v)) - \chi_{a,m}(F_{W|A=m}(v))}{F_{W|A=m,T}(v) - F_{W|A=m}(v)} - \chi'_{a,m}(F_{W|A=m}(v))dv \right) dv \right| < \frac{\varepsilon}{2} \tag{13}$$

when T is large enough. Because of the properties of $\chi_{a,m}$ in (ii) and (iii) in Lemma W1 and the assumption that the distribution $F_{W|A=m}$ is strictly increasing, we can pick a constant η^* small enough so that (a) $\chi_{a,m}$ is convex over the interval between $F_{W|A=m}(\overline{v} - \eta^*)$ and 1; and (b) the ratio $\frac{1-\chi_{a,m}(F_{W|A=m}(v))}{\sqrt{1-F_{W|A=m}(v)}}$ is bounded above by a finite constant uniformly over $v \in [\overline{v} - \eta^*, \overline{v}]$. This implies

$$\begin{aligned} 0 < \int_{\overline{v}-\eta^*}^{\overline{v}} \chi'_{a,m}(F_{W|A=m}(v))dv &< \int_{\overline{v}-\eta^*}^{\overline{v}} \frac{\chi_{a,m}(F_{W|A=m}(\overline{v})) - \chi_{a,m}(F_{W|A=m}(v))}{F_{W|A=m}(\overline{v}) - F_{W|A=m}(v)} dv \\ = \int_{\overline{v}-\eta^*}^{\overline{v}} \frac{1 - \chi_{a,m}(F_{W|A=m}(v))}{\sqrt{1 - F_{W|A=m}(v)}} \frac{1}{\sqrt{1 - F_{W|A=m}(v)}} dv &\leq \mathcal{C}_{a,m,\eta^*} \left(\int_{\overline{v}-\eta^*}^{\overline{v}} \frac{1}{\sqrt{1 - F_{W|A=m}(v)}} dv \right) \end{aligned}$$

where the second (strict) inequality is due to the convexity of $\chi_{a,m}$ over $[F_{W|A=m}(\overline{v} - \eta^*), 1]$; the third (weak) inequality is due to the consequence (b) of the choice of η^* above. The constant $\mathcal{C}_{a,m,\eta^*}$ depends on a, m and is decreasing in the choice of η^*. Thus it follows from Lemma W2 that, for any $\varepsilon > 0$, we can pick $\eta_1(\varepsilon)$ small enough so that $\int_{\overline{v}-\eta_1(\varepsilon)}^{\overline{v}} \chi'_{a,m}(F_{W|A=m}(v))dv \in (0, \varepsilon/2)$. Lemma W3 implies we can choose $\eta_2(\varepsilon)$ so that for any $F_{W|A=m,T} \in \hat{\mathcal{D}}_{[\underline{v},\overline{v}]}$ with $\sqrt{T}\left(F_{W|A=m,T} - F_{W|A=m}\right) \to F^*$ for some $F^* \in \mathcal{B}_{[\underline{v},\overline{v}]}$, $\int_{\overline{v}-\eta_2(\varepsilon)}^{\overline{v}} \left(\frac{\chi_{a,m}(F_{W|A=m,T}(v)) - \chi_{a,m}(F_{W|A=m}(v))}{F_{W|A=m,T}(v) - F_{W|A=m}(v)} \right) dv \in (0, \varepsilon/2)$ when T is sufficiently large. Let $\delta(\varepsilon) \equiv \min\{\eta_1(\varepsilon), \eta_2(\varepsilon)\}$. Then (13) holds for any such $F_{W|A=m,T} \in \hat{\mathcal{D}}_{[r,\overline{v}]}$ when T is sufficiently large.

Next, note the ratio between $\chi_{a,m}(F_{W|A=m,T}(v)) - \chi_{a,m}(F_{W|A=m}(v))$ and $F_{W|A=m,T}(v) - F_{W|A=m}(v)$ converges to $\chi'_{a,m}(F_{W|A=m}(v))$ pointwise at any $v \in [r, \overline{v} - \delta(\varepsilon)]$. Since $\overline{v} - \delta(\varepsilon) < \overline{v}$ and $F_{W|A=m}$ is increasing, $F(\overline{v} - \delta(\varepsilon))$ is strictly less than 1. Hence for any sequence $F_{W|A=m,T} \in \hat{\mathcal{D}}_{[\underline{v},\overline{v}]}$ such that $\sqrt{T}(F_{W|A=m,T} - F_{W|A=m}) \to F^*$ for some $F^* \in \mathcal{B}_{[\underline{v},\overline{v}]}$, the ratio between $\chi_{a,m}(F_{W|A=m,T}(v)) - \chi_{a,m}(F_{W|A=m}(v))$ and $F_{W|A=m,T}(v) - F_{W|A=m}(v)$ is bounded above by a finite positive constant uniformly over $[r, \overline{v} - \delta(\varepsilon)]$ when T is large enough. It then follows from the Dominated Convergence theorem that

$$\lim_{T\to\infty} \int_r^{\overline{v}-\delta(\varepsilon)} \left(\frac{\chi_{a,m}(F_{W|A=m,T}(v)) - \chi_{a,m}(F_{W|A=m}(v))}{F_{W|A=m,T}(v) - F_{W|A=m}(v)} \right) dv = \int_r^{\overline{v}-\delta(\varepsilon)} \chi'_{a,m}(F_{W|A=m}(v))dv$$

for any such sequence $F_{W|A=m,T} \in \hat{\mathcal{D}}_{[\underline{v},\overline{v}]}$. Or equivalently, for any such sequence $F_{W|A=m,T} \in \hat{\mathcal{D}}_{[\underline{v},\overline{v}]}$,

$$\left| \int_r^{\overline{v}-\delta(\varepsilon)} \left(\frac{\chi_{a,m}(F_{W|A=m,T}(v)) - \chi_{a,m}(F_{W|A=m}(v))}{F_{W|A=m,T}(v) - F_{W|A=m}(v)} - \chi'_{a,m}(F_{W|A=m}(v))dv \right) dv \right| < \frac{\varepsilon}{2} \tag{14}$$

when T is large enough. Combining the two inequalities from (13) and (14) then proves the lemma. $\square$

Proof of Lemma B4. The smoothness of $F_{W|A=m}$ due to Lemma W2 implies that for any sequence $F_{W|A=m,T} \in \hat{\mathcal{D}}_{[\underline{v},\overline{v}]}$, the set $\{v \in [\underline{v},\overline{v}] : F_{W|A=m,T}(v) = F_{W|A=m}(v)\}$ has zero Lebesgue measure. Thus for all T,

$$\begin{aligned} & \int_r^{\overline{v}} \sqrt{T}\left[\chi_{a,m}(F_{W|A=m,T}(v)) - \chi_{a,m}(F_{W|A=m}(v))\right] dv \qquad (15) \\ = \; & \int_r^{\overline{v}} \sqrt{T}[F_{W|A=m,T}(v) - F_{W|A=m}(v)] \left(\frac{\chi_{a,m}(F_{W|A=m,T}(v)) - \chi_{a,m}(F_{W|A=m}(v))}{F_{W|A=m,T}(v) - F_{W|A=m}(v)}\right) dv. \end{aligned}$$

For any sequence $F_{W|A=m,T} \in \hat{\mathcal{D}}_{[\underline{v},\overline{v}]}$ such that $\sqrt{T}(F_{W|A=m,T} - F_{W|A=m}) \to F^*$ for some $F^* \in \mathcal{B}_{[\underline{v},\overline{v}]}$ and for any constant $\varepsilon^* > 0$, $\sqrt{T}(F_{W|A=m,T} - F_{W|A=m})$ must be bounded between $F^* - \varepsilon^*$ and $F^* + \varepsilon^*$ uniformly over $[r,\overline{v}]$ when T is sufficiently large. This means for T large enough, the absolute value of the integrand on the right-hand side of (15) is bounded above by the product of a finite, positive constant and the integral $\int_r^{\overline{v}} \left(\frac{\chi_{a,m}(F_{W|A=m,T}(v)) - \chi_{a,m}(F_{W|A=m}(v))}{F_{W|A=m,T}(v) - F_{W|A=m}(v)}\right) dv$, which is positive due to the monotonicity of $\chi_{a,m}$. By the Dominated Convergence Theorem in its general form (see, for example, Theorem 2.2.2-(iii) in Lehmann and Romano (2005)), Lemma B4 would hold if we can show

$$\lim_{T\to\infty} \int_r^{\overline{v}} \left(\frac{\chi_{a,m}(F_{W|A=m,T}(v)) - \chi_{a,m}(F_{W|A=m}(v))}{F_{W|A=m,T}(v) - F_{W|A=m}(v)}\right) dv = \int_r^{\overline{v}} \chi'_{a,m}(F_{W|A=m}(v)) dv$$

for all sequences $F_{W|A=m,T} \in \hat{\mathcal{D}}_{[\underline{v},\overline{v}]}$ such that $\sqrt{T}\left(F_{W|A=m,T} - F_{W|A=m}\right) \to F^*$ for some $F^* \in \mathcal{B}_{[\underline{v},\overline{v}]}$. This is shown above in Lemma W4. $\square$

3 The Covariance Formula for Lemma B5

Let $s < v$ on $[\underline{v},\overline{v}]$. To simplify notation, we suppress dependence on the sample size T throughout the note. By Assumptions 1 and 2 and the multivariate version of Lindeberg-Levy Central Limit Theorem (CLT),

$$\sqrt{T}\begin{pmatrix} \frac{1}{T}\sum_{t\le T} 1\{W_t \le s, A_t = m\} - p_{sm} \\ \frac{1}{T}\sum_{t\le T} 1\{W_t \le v, A_t = m\} - p_{vm} \\ \frac{1}{T}\sum_{t\le T} 1\{A_t = m\} - p_m \\ \frac{1}{T}\sum_{t\le T} \tilde{K}_t - \mu_K \end{pmatrix} \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \mathbf{\Sigma}_0\right) \qquad (16)$$

where

$$\mathbf{\Sigma}_0 \equiv \begin{bmatrix} p_{sm}(1-p_{sm}) & p_{sm}(1-p_{vm}) & p_{sm}(1-p_m) & \mu_{sm} - p_{sm}\mu_K \\ p_{sm}(1-p_{vm}) & p_{vm}(1-p_{vm}) & p_{vm}(1-p_m) & \mu_{vm} - p_{vm}\mu_K \\ p_{sm}(1-p_m) & p_{vm}(1-p_m) & p_m(1-p_m) & \mu_m - p_m\mu_K \\ \mu_{sm} - p_{sm}\mu_K & \mu_{vm} - p_{vm}\mu_K & \mu_m - p_m\mu_K & \sigma_K^2 + \sigma_\epsilon^2 \end{bmatrix}$$

with $p_{sm} \equiv \Pr(W \le s, A = m)$, $p_m \equiv \Pr(A = m)$, $\mu_m \equiv E[1\{A = m\}K]$ and $\mu_{sm} \equiv E[1\{W \le s, A = m\}K]$. Note we have used the independence between ϵ and K and that $E(\epsilon) = 0$. We use (16) as a building block for deriving the covariance kernel between $\mathbb{F}_{W|m}$, $\mathcal{N}_\rho$ and $\mathcal{N}_\mu$.

First, by the Delta Method,

$$\sqrt{T}\begin{pmatrix} \hat{F}_{W|A=m}(s) - F_{W|A=m}(s) \\ \hat{F}_{W|A=m}(v) - F_{W|A=m}(v) \end{pmatrix} \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \mathbf{\Sigma}_1\right)$$

where

$$\boldsymbol{\Sigma}_1 = \begin{pmatrix} \frac{1}{p_m} & 0 & -\frac{p_{sm}}{p_m^2} \\ 0 & \frac{1}{p_m} & -\frac{p_{vm}}{p_m^2} \end{pmatrix} \begin{pmatrix} p_{sm}(1-p_{sm}) & p_{sm}(1-p_{vm}) & p_{sm}(1-p_m) \\ p_{sm}(1-p_{vm}) & p_{vm}(1-p_{vm}) & p_{vm}(1-p_m) \\ p_{sm}(1-p_m) & p_{vm}(1-p_m) & p_m(1-p_m) \end{pmatrix} \begin{pmatrix} \frac{1}{p_m} & 0 \\ 0 & \frac{1}{p_m} \\ -\frac{p_{sm}}{p_m^2} & -\frac{p_{vm}}{p_m^2} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{p_m^3} p_{sm}(p_m - p_{sm}) & \frac{1}{p_m^3} p_{sm}(p_m - p_{vm}) \\ \frac{1}{p_m^3} p_{sm}(p_m - p_{vm}) & \frac{1}{p_m^3} p_{vm}(p_m - p_{vm}) \end{pmatrix}.$$

Second, by construction, for any $a < b$,

$$\sqrt{T}\begin{pmatrix} \hat{\rho}_{b,T} - \rho_b \\ \hat{\rho}_{a,T} - \rho_a \end{pmatrix} = \sqrt{T}\begin{pmatrix} \frac{1}{T}\sum_{t\le T}\left[\frac{N-b}{N}1\{A_t = b\} + \frac{b+1}{N}1\{A_t = b+1\}\right] - \left(\frac{N-b}{N}p_b + \frac{b+1}{N}p_{b+1}\right) \\ \frac{1}{T}\sum_{t\le T}\left[\frac{N-a}{N}1\{A_t = a\} + \frac{a+1}{N}1\{A_t = a+1\}\right] - \left(\frac{N-a}{N}p_a + \frac{a+1}{N}p_{a+1}\right) \end{pmatrix}.$$

It is straightforward to show by the multivariate CLT that

$$\sqrt{T}\begin{pmatrix} \hat{\rho}_{b,T} - \rho_b \\ \hat{\rho}_{a,T} - \rho_a \end{pmatrix} \xrightarrow{d} \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_2)$$

where

$$\boldsymbol{\Sigma}_2 = \begin{pmatrix} (\frac{N-b}{N})^2 p_b + (\frac{b+1}{N})^2 p_{b+1} - \rho_b^2 & \bar{\rho}_{a,b} - \rho_a\rho_b \\ \bar{\rho}_{a,b} - \rho_a\rho_b & (\frac{N-a}{N})^2 p_a + (\frac{a+1}{N})^2 p_{a+1} - \rho_a^2 \end{pmatrix}$$

with $\bar{\rho}_{a,b} \equiv \frac{(N-b)b}{N^2}p_b$ if $b = a+1$; and $\bar{\rho}_{a,b} \equiv 0$ otherwise. Also by the multivariate CLT,

$$\sqrt{T}\begin{pmatrix} \hat{\rho}_{a,T} - \rho_a \\ \frac{1}{T}\sum_t \tilde{K}_t - \mu_K \end{pmatrix} \xrightarrow{d} \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_3)$$

where

$$\boldsymbol{\Sigma}_3 = \begin{pmatrix} (\frac{N-a}{N})^2 p_a + (\frac{a+1}{N})^2 p_{a+1} - \rho_a^2 & \frac{N-a}{N}\mu_a + \frac{a+1}{N}\mu_{a+1} - \rho_a\mu_K \\ \frac{N-a}{N}\mu_a + \frac{a+1}{N}\mu_{a+1} - \rho_a\mu_K & \sigma_K^2 + \sigma_\epsilon^2 \end{pmatrix}.$$

Note we have used the independence between ϵ and K and the zero mean of ϵ here.

Third,

$$\sqrt{T}\begin{pmatrix} \hat{F}_{W|m}(s) - F_{W|m}(s) \\ \frac{1}{T}\sum_t \tilde{K}_t - \mu_K \end{pmatrix} \xrightarrow{d} \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_4)$$

where

$$\boldsymbol{\Sigma}_4 = \begin{pmatrix} \frac{1}{p_m} & -\frac{p_{sm}}{p_m^2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} p_{sm}(1-p_{sm}) & p_{sm}(1-p_m) & \mu_{sm} - p_{sm}\mu_K \\ p_{sm}(1-p_m) & p_m(1-p_m) & \mu_m - p_m\mu_K \\ \mu_{sm} - p_{sm}\mu_K & \mu_m - p_m\mu_K & \sigma_K^2 + \sigma_\epsilon^2 \end{pmatrix} \begin{pmatrix} \frac{1}{p_m} & 0 \\ -\frac{p_{sm}}{p_m^2} & 0 \\ 0 & 1 \end{pmatrix}$$

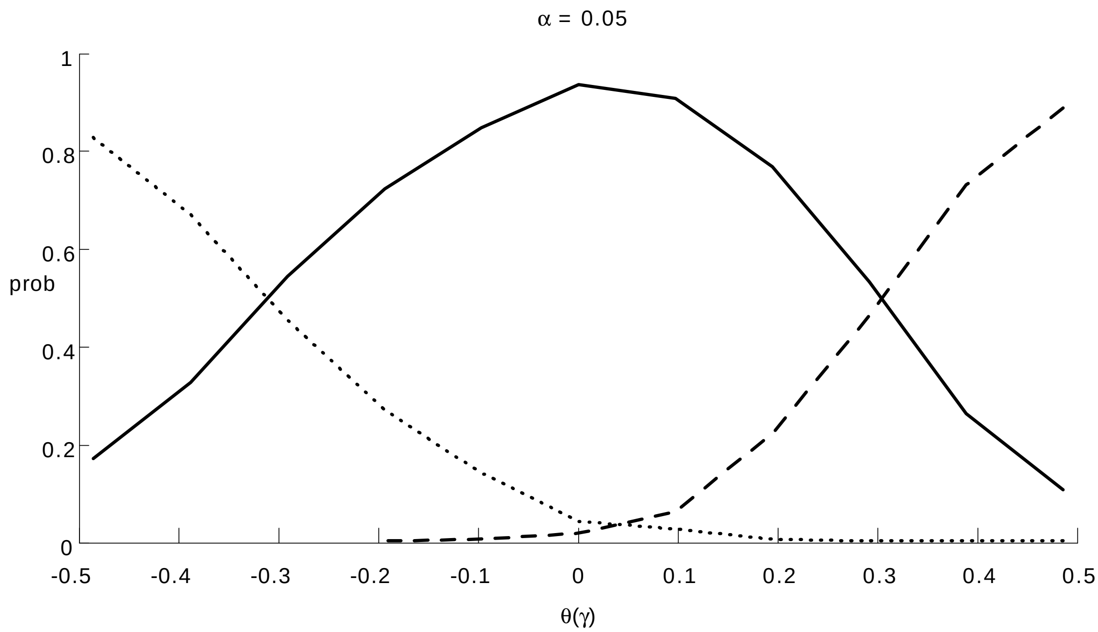
$$= \begin{pmatrix} \frac{1}{p_m^3} p_{sm}(p_m - p_{sm}) & \frac{1}{p_m^2}(p_m\mu_{sm} - \mu_m p_{sm}) \\ \frac{1}{p_m^2}(p_m\mu_{sm} - \mu_m p_{sm}) & \sigma_K^2 + \sigma_\epsilon^2 \end{pmatrix}.$$

Fourth,

$$\sqrt{T}\begin{pmatrix} \hat{F}_{W|m}(s) - F_{W|m}(s) \\ \hat{\rho}_a - \rho_a \end{pmatrix} \xrightarrow{d} \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_5)$$

where $\boldsymbol{\Sigma}_5 = \tilde{\mathbf{D}}\tilde{\boldsymbol{\Sigma}}\left(\tilde{\mathbf{D}}\right)'$, and $\tilde{\boldsymbol{\Sigma}}$ is the covariance matrix of $[1\{W \le s, A = m\}, 1\{A = m\}, 1\{A = a\}, 1\{A = a+1\}]'$ if $m > a+1$ or $m < a$; and is the covariance matrix of $[1\{W \le s, A = m\}, 1\{A = a\}, 1\{A = a+1\}]'$ otherwise; and $\tilde{\mathbf{D}}$ is the 2-by-4 (or 2-by-3) Jacobian matrix needed to apply the multivariate Delta Method. Under Assumptions 1 and 2, both $\tilde{\mathbf{D}}$ and $\tilde{\boldsymbol{\Sigma}}$ have full rank.

Appendix : Figures



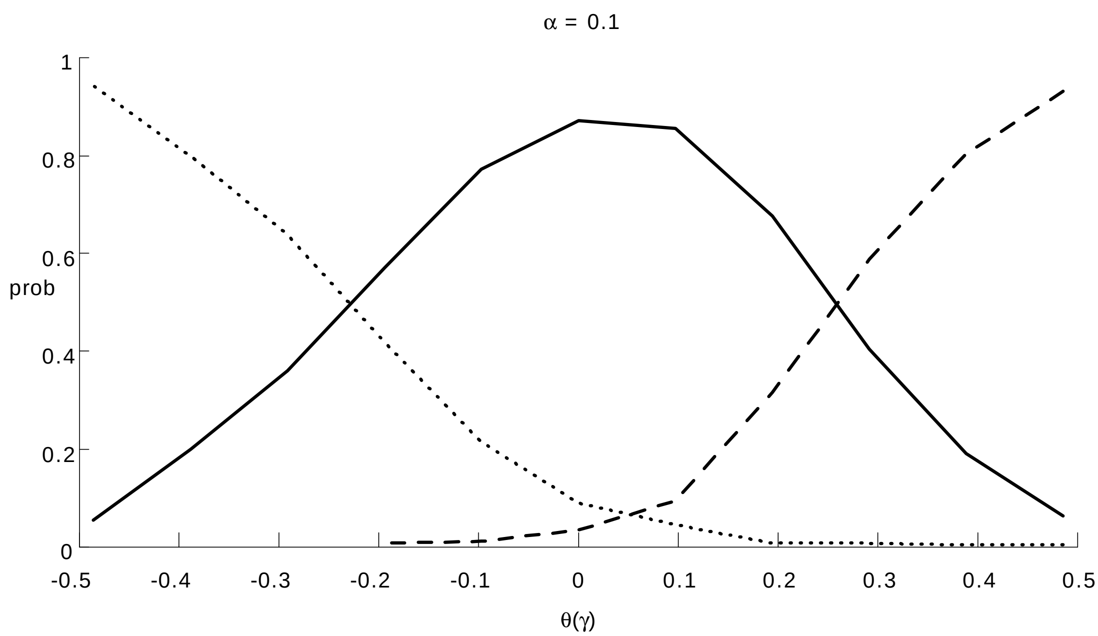


Figure 1 (a): Test Performance ($N = 4$, $T = 1500$)

Notes: Horizontal axis: Integrated Measure of Risk Aversion. Solid line: proportion of $S = 250$ simulated samples in which our test fails to reject the null (risk-neutrality). Dashed line: proportion that the null is rejected in favor of H_A (risk-aversion). Dotted line: proportion that the null is rejected in favor of H_L (risk-loving).

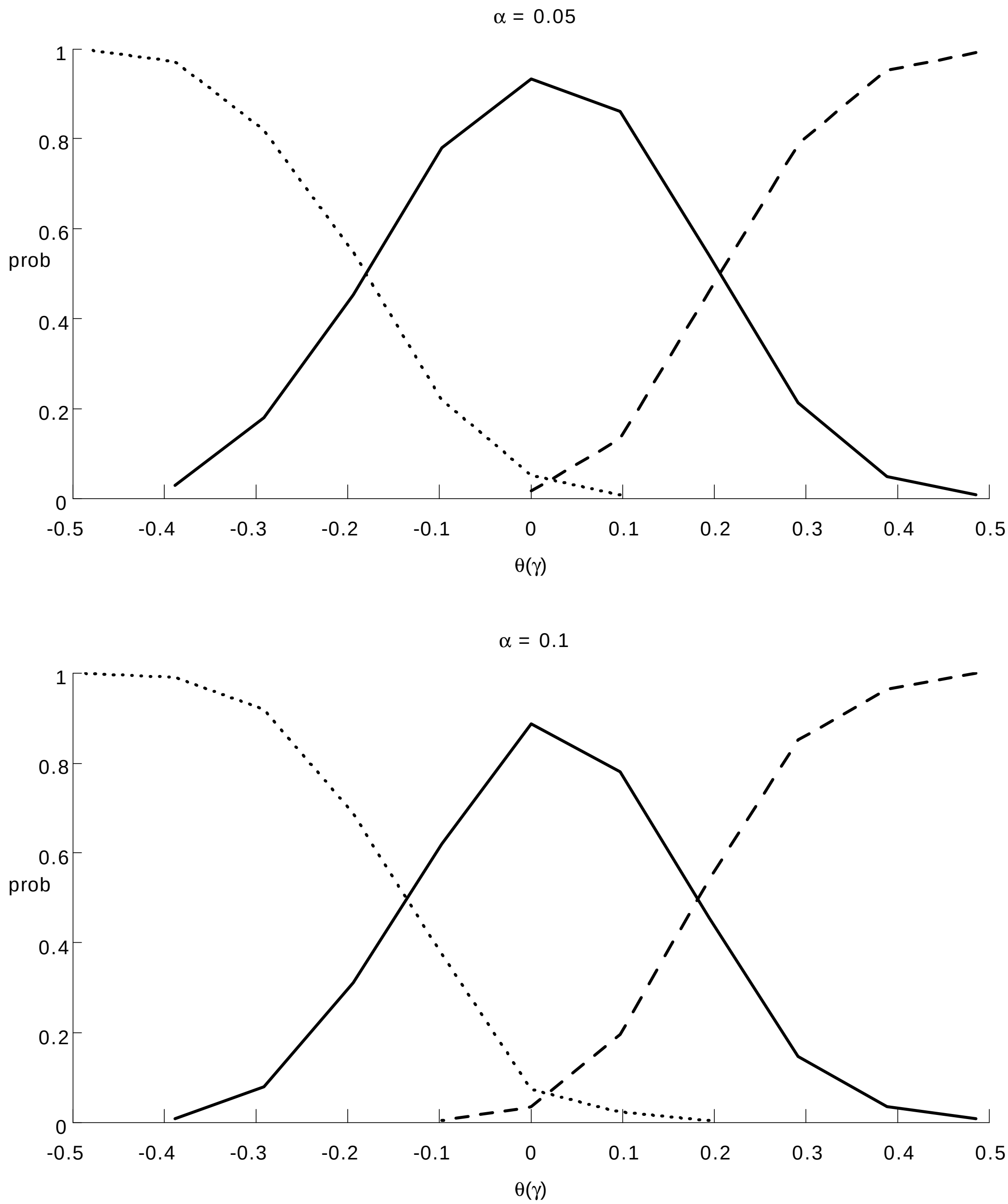


Figure 1 (b): Test Performance $(N = 4,\, T = 3000)$

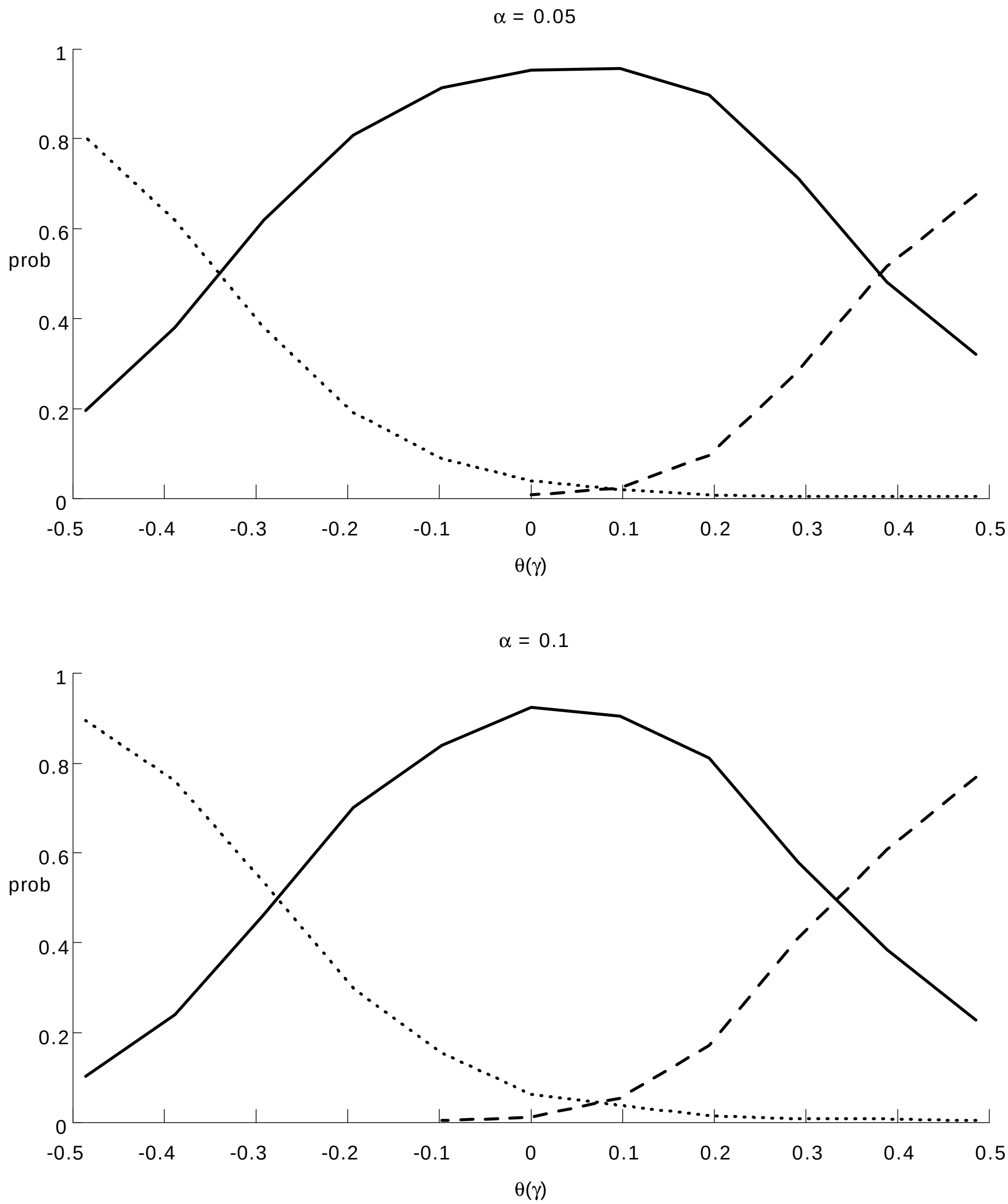


Figure 1 (c): Test Performance $(N = 5,\, T = 1500)$

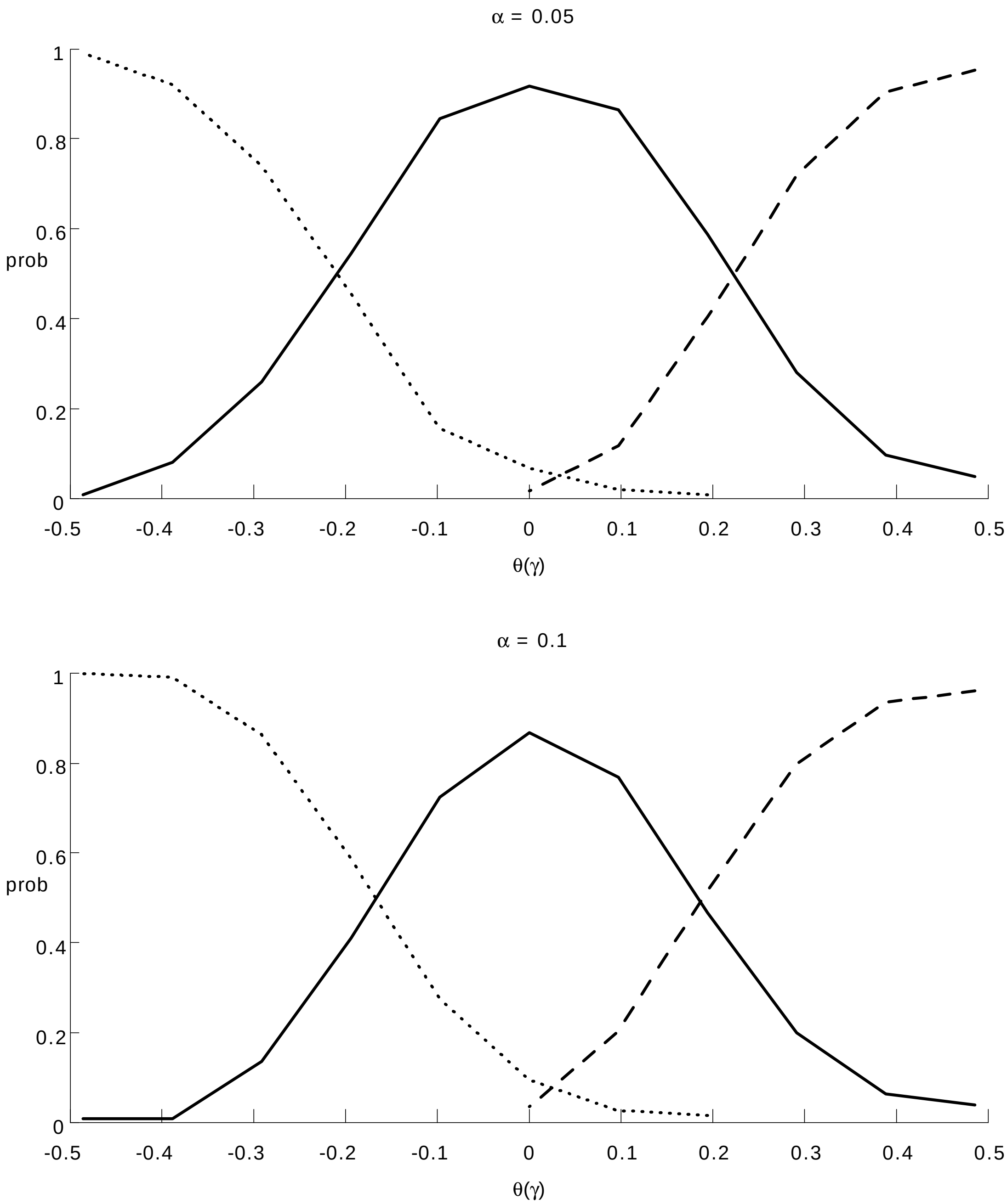


Figure 1 (d): Test Performance $(N = 5,\, T = 3000)$

REFERENCE

Fang, H. and X. Tang (2014). "Inference of Bidders' Risk Attitudes in Ascending Auctions with Endogenous Entry," unpublished manuscript, University of Pennsylvania.

Kosorok, M. (2008). *Introduction to Empirical Processes and Semiparametric Inference*. Springer-Verlag: New York.

Levin, D. and J. Smith (1994). "Equilibrium in Auctions with Entry," *American Economic Review*, Vol. 84, No. 3, 585-599.

Menzel, K. and P. Morganti. (2013). "Large sample properties for estimators based on the order statistics approach in auctions," *Quantitative Economics*, 4: 329–375.

Van der Vaart, A. and J. Wellner (1996). *Weak Convergence and Empirical Processes*. Springer-Verlag: New York.