

Penn Institute for Economic Research  
Department of Economics  
University of Pennsylvania  
3718 Locust Walk  
Philadelphia, PA 19104-6297  
[pier@econ.upenn.edu](mailto:pier@econ.upenn.edu)  
<http://economics.sas.upenn.edu/pier>

## *PIER Working Paper 13-056*

“Inference of Bidders’ Risk Attitudes in Ascending Auctions with Endogenous Entry”

by

Hanming Fang and Xun Tang

<http://ssrn.com/abstract=2339219>

# Inference of Bidders' Risk Attitudes in Ascending Auctions with Endogenous Entry\*

Hanming Fang<sup>†</sup>      Xun Tang<sup>‡</sup>

September 9, 2013

## Abstract

Bidders' risk attitudes have key implications for choices of revenue-maximizing auction formats. In ascending auctions, bid distributions do not provide information about risk preference. We infer risk attitudes using distributions of transaction prices and participation decisions in ascending auctions with entry costs. Nonparametric tests are proposed for two distinct scenarios: first, the expected entry cost can be consistently estimated from data; second, the data does not report entry costs but contains exogenous variations of potential competition and auction characteristics. In the first scenario, we exploit the fact that the risk premium required for entry – the difference between *ex ante* expected profits from entry and the certainty equivalent – is strictly positive if and only if bidders are risk averse. Our test is based on identification of bidders' *ex ante* profits. In the second scenario, our test builds on the fact that risk attitudes affect how equilibrium entry probabilities vary with observed auction characteristics and potential competition. We also show identification of risk attitudes in a more general model of ascending auctions with selective entry, where bidders receive entry-stage signals that are correlated with private values.

**Keywords:** Ascending auctions, Risk attitudes, Endogenous entry, Tests

**JEL Classification Codes:** D44, C12, C14

---

\*We are grateful to the associate editor and three anonymous referees whose constructive comments substantially improved the paper. We also thank Dan Akerberg, Federico Bugni, Xu Cheng, Flavio Cunha, Ken Hendricks, Tong Li, Isabelle Perrigne, Frank Schorfheide, Steven Stern, Petra Todd, Quang Vuong and seminar/conference participants at UC Berkeley, Michigan, Northwestern, Pennsylvania, Toronto, Virginia, Wharton, North American Econometric Society Summer Meetings (St. Louis 2011) and Society of Economic Dynamics Annual Meetings (Ghent 2011) for helpful comments and suggestions. All remaining errors are our own.

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

<sup>‡</sup>Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104. Email: xuntang@econ.upenn.edu

# 1 Introduction

We propose nonparametric tests to infer bidders' risk attitudes in ascending (or open out-cry) auctions with endogenous entry. In these auctions, potential bidders observe some entry costs, e.g., bid preparation/submission costs and/or information acquisition costs, that they need to incur before learning private values, and decide whether to pay the costs to be active in the bidding stage. Bidders make rational entry decisions by comparing expected utility from entry with that from staying out, based on their knowledge of entry costs or preliminary signals of private values to be realized in the subsequent bidding stage.

Inference of bidders' risk attitudes have important implications for sellers' choices of revenue-maximizing auction format. When participation of bidders is exogenously given and fixed, the Revenue Equivalence Theorem states that expected revenues from first-price and ascending auctions are the same if bidders are risk-neutral in symmetric independent private values (IPV) environments. If bidders are risk-averse, however, Matthews (1987) showed that in such environments first-price auctions yield higher expected revenues than ascending auctions. Bidders' risk attitudes also affect revenue rankings among symmetric IPV auctions when participation decisions are endogenous. For risk-neutral bidders, Levin and Smith (1994) showed that any given entry cost induces the same entry probabilities in first-price auctions (with entrants observing the number of other entrants) and in ascending auctions. Thus the Revenue Equivalence Theorem implies expected revenues must be the same from both first-price and ascending formats under endogenous entry. On the other hand, Smith and Levin (1996) established the revenue ranking of first-price over ascending auctions under endogenous entry for risk-averse bidders, except for the case with decreasing absolute risk aversions (DARA).<sup>1</sup>

While some earlier papers had studied the identification and estimation of bidders' risk attitudes in *first-price* auctions (e.g. Bajari and Hortascu (2005), Campo, Guerre, Perrigne and Vuong (2011) and Guerre, Perrigne and Vuong (2009)), inference of risk attitudes in ascending auctions remains an open question. Athey and Haile (2007) pointed out that bidders' risk attitudes cannot be identified from bids *alone* in ascending auctions where participation is given exogenously. This is because bidding one's true value is a weakly dominant strategy in ascending auctions, *regardless of bidders' risk attitudes*. Thus, bidders with various risk attitudes could generate the same distribution of bids in Bayesian Nash

---

<sup>1</sup>Even in the case with DARA, first-price auctions yield higher expected revenues than ascending auctions when entry costs are low enough. To see this, consider a simple case where entry costs are low enough so that the difference between entry probabilities in first-price and ascending auctions are sufficiently small. In such a case, these two probabilities are both close to 1 and only differ by some small  $\varepsilon > 0$ . By Matthews (1987), conditioning on any given number of entrants, ascending auctions have smaller expected revenues than first-price auctions. With difference between the two entry probabilities  $\varepsilon$  small enough, such a revenue ranking result will be preserved.

equilibria. The distribution of bids from entrants alone is therefore not sufficient for inferring bidders' risk attitudes.

In this paper, we propose tests for bidders' risk-attitudes based on transaction prices *and* entry decisions under two empirically relevant data scenarios. In the first scenario, we assume that the data has information that allows researchers to consistently estimate the expectation of entry costs; and in the second scenario, we assume that the data does not provide any information about the level of entry costs, but does contain variations in the number of potential bidders and some auction-level heterogeneities. In both cases, we show how to relate the distribution of transaction prices and entry decisions to the underlying risk attitudes nonparametrically.

In the first scenario, considered in Section 4, we require that the data contains some noisy measures of entry costs so that the mean of entry costs in the data-generating process can be consistently estimated. This is motivated by the fact that entry costs are often measurable (at least up to some noises) in applications. Examples of entry costs include bid preparation costs (e.g., mailing costs), admission fees or other information acquisition expenses, which are often reported in data with noises.

The main insight for our test in this scenario can be illustrated using the mixed-strategy entry model (which is analogous to that considered in Levin and Smith (1994) for first-price auctions). In the entry stage, all potential bidders observe some common entry cost and decide whether to pay the cost and enter an ascending auction in the bidding stage. In a mixed strategy Nash equilibrium in the entry stage, potential bidders' participation in the auction will be in mixed strategies with the mixing probability determined to ensure that a bidder's expected *utility* from entry equals that from staying out. Hence bidders' risk attitudes can be identified by comparing the expected *profits* from entry and its certainty equivalent. As long as the expectation of entry costs can be identified from data, the distribution of transaction prices and entry decisions alone can be used to make such a comparison.

We apply the analog principle to construct a non-parametric test statistic, using data on transaction prices and entry decisions as well as estimates of the mean of entry costs. We characterize the limiting distribution of this statistic, and propose a bootstrap test that attains correct asymptotic level and is consistent under any fixed alternative of risk-aversion or risk-loving. We provide evidence for its finite sample performance through Monte Carlo simulations.

In the second scenario, considered in Section 5, we assume that the data does not provide information about the level of entry costs, but does contain some auction-level heterogeneities (such as feature of the auctioned object) and variations in the numbers of potential bidders. We propose a nonparametric test for risk-aversion for this case, under the assumption that variations in potential competition are exogenous (in the sense that it does not alter the marginal distribution of private values once conditional on the observed auction features).

Refraining from parametric restrictions on how observed auction features change the distribution of private values, we formalize how risk attitudes determine entry probabilities under various auction features and potential competition. Our test is based on the idea that the curvature of utility functions affects how the ratios between the changes in bidders’ interim utilities compare with the ratios between the changes in expected private values, under different pairs of observed auction features. The main finding is that, even if entry costs are unreported in data, we could relate these two ratios to the observed distributions of transaction prices and entry decisions, by exploiting the variations in auction features and potential number of entrants.

In Section 6, we discuss possible extensions of our tests by removing two of the key assumptions. First, we show that when bidders’ values are affiliated, it is possible to derive testable implications for risk attitudes using existing results on sharp bounds for the surplus of risk-neutral bidders in ascending auctions (Aradillas-Lopez, Gandhi and Quint (2012)). Second, we show that if entry becomes selective (e.g., due to the presence of preliminary signals correlated with private values to be drawn in the bidding stage), then the idea of testing risk attitudes through identification of risk-premium applies, provided the data contains continuous variations in observed entry costs.

It is worth noting that by “inference of risk attitudes” we mean to make a data-supported conclusion about whether bidders’ utility functions are concave, linear or convex. We do not deal with the recovery of the full utility functions per se, which is a question left for future research.

The remainder of the paper is structured as follows. In Section 2 we discuss the related literature; in Section 3 we present the model of ascending auction with endogenous entry; in Sections 4 and 5 we describe the theoretical results for our tests under two data scenarios, and discuss inference using proposed test statistics; in Section 6 we discuss how to extend our test to auctions with selective entry or affiliated private values; and in Section 7 we conclude. Proofs are collected in the appendices.

## 2 Related Literature

This paper contributes to two branches of the literature on structural analyses of auction data. The first branch includes papers that analyze the equilibrium and its empirical implications in auctions with endogenous entry and *risk-neutral* bidders. These include Levin and Smith (1994), Li (2005), Ye (2007) and Li and Zheng (2009). Marmer, Shneyerov and Xu (2011) studies a model of first-price auctions with risk-neutral bidders and selective entry, and discusses testable implications of various nested entry models. Roberts and Sweeting (2010) estimates a model of ascending auctions with selective entry and risk-neutral bidders. Gentry and Li (2013) provides partial identification results for ascending auctions with

risk-neutral bidders when entry is selective (as potential bidders observe preliminary signals that are affiliated with their private values). They derive sharp bounds on the distribution of private values conditional on signals, using variations in factors that affect bidders' entry behaviors (such as the number of potential bidders and entry costs). They also apply these results to bound counterfactual seller revenues under alternative auction rules. Aradillas-Lopez, Gandhi and Quint (2012) provides partial identification results for ascending auctions where bidders' private values are affiliated, exploiting exogenous variations in the number of entrants, or active bidders.

The second branch includes papers that study the identification and estimation of bidders' utility functions and the distribution of private values in first-price auctions without endogenous entry. Campo, Guerre, Perrigne and Vuong (2011) shows how to estimate a semiparametric model of *first-price* auctions with risk-averse bidders when the identification of a parametric utility function is assumed. Bajari and Hortascu (2005) uses exogenous variations in the number of bidders in first-price auctions to semi-parametrically estimate the utility function while leaving the distribution of bidders' private values unrestricted. Guerre, Perrigne and Vuong (2009) uses exogenous variations in the number of potential bidders to non-parametrically identify bidders' utility functions along with the distribution of private values in first-price auctions. Lu and Perrigne (2008) considers a context where data contain bids from both first-price and ascending auctions that involve bidders with the same utility function and the same distribution of private values. They first use bids from ascending auctions to recover the distribution of private values, and then use bids from first-price auctions to recover the utility function.

Our work in this paper contributes to these two branches of empirical auction literature by studying a model which endogenizes bidders' entry decisions and relaxes the risk-neutrality assumption at the same time. To the best of our knowledge, our paper marks the first effort to non-parametrically infer bidders' risk attitudes in ascending auctions with endogenous entry. Smith and Levin (1996) presents some results on the revenue ranking of auction formats in terms of seller revenues when auctions are known to involve risk-averse bidders who make endogenous entry decisions. Their focus is not on the identification of bidders' risk attitudes.

Ackerberg, Hirano and Shahriar (2011) studies a class of e-Bay auctions where a typical online ascending auction is combined with an option of paying the buy-out price posted by the seller in order to immediately purchase the object. They show how to identify the bidders' utility functions and the distribution of private values using exogenous variations in the buy-out prices and other auction characteristics. The format of auctions they consider is qualitatively different from the one we consider in this paper, which is a standard ascending format with endogenous entry. We do not embark on a full identification of the utility function in this paper, and therefore require fewer sources of exogenous variations to perform

the test. (With observed and exogenous variations in entry costs, identification of the utility function may be possible in our model as well.) Our approach does not rely on variations in entry costs, thus our test can be performed for any given level of entry costs. Moreover, we go beyond identification and propose a method of robust inference.

Our paper also fits in a category of empirical auction literature on nonparametric tests of the empirical implications/predictions of auction theory. Earlier works in this category include tests of bidders' rationality in first-price auctions with common values in Hendricks, Pinkse and Porter (2003), assessment of winner's curse in first-price auctions with common values in Hong and Shum (2002), tests for presence of interdependent values in Haile, Hong and Shum (2004), and test for affiliations between bidders' private values in Li and Zhang (2010) and Jun, Pinkse and Wan (2010).<sup>2</sup> The test proposed by Li and Zhang (2010) uses bidders' entry decisions for testing affiliations between bidders' private values.

Li, Lu and Zhao (2012) studies first-price and ascending auctions with risk-averse bidders and selective entry (due to bidders' observation of preliminary signals that are affiliated with private values in the bidding stage). They find the ranking between bidders' expected utility under these two auction formats depend on the specific form of risk-aversion (DARA, CARA or IARA). Consequently, the ranking of entry probabilities across these two auction formats must also depend on the form of risk-aversions. Based on these observations, they propose an original test for inferring the form of bidders' risk-aversions using entry behaviors, using data from both first-price and ascending auctions at the same time. Our research focus in this paper differs from theirs in that we study the inference of risk attitudes (with hypotheses being risk-neutral, risk-taking or risk-aversion). Since we do not aim at inferring the form of risk-aversion, our test does not require observation of data from competing auction formats.

### 3 Ascending Auctions with Endogenous Entry

Consider an empirical context where researchers observe data from a large number of independent single-unit ascending auctions. Each of these auctions involve  $N$  potential bidders who have symmetric independent private values and make endogenous entry decisions. The entry model we consider here is similar to those used for first-price auctions with risk-neutral bidders in Levin and Smith (1994), Li and Zheng (2009) and Marmer, Shneyerov and Xu (2011). In the entry stage, each potential bidder decides whether to incur an entry cost  $K$  to become active in the bidding stage. The entry cost is common knowledge among all potential bidders. There is a binding reserve price  $r$  in the auction and it is observed by all potential bidders in the entry stage. Following their entry, each entrant  $i$  draws a private value  $V_i$ , and competes in ascending auctions in the bidding stage. Across auctions, private values and entry costs are independent draws from distributions  $F_{V_1, \dots, V_N | K}$  and  $F_K$  respectively, which

---

<sup>2</sup>See Athey and Haile (2007) and Hendricks and Porter (2007) for recent surveys.

are common knowledge among all potential bidders prior to entry decisions. Upon entry, each entrant may or may not be aware of the total number of entrants (which we denote by  $A$ ). All bidders in the data share the same bounded von Neumann-Morgenstein utility function  $u : \mathbb{R}_+ \rightarrow \mathbb{R}$  with  $u' > 0$  and the sign of  $u''$  does not change over  $\mathbb{R}_+$ . A winner who has a private value  $V_i$  and pays a price  $P_i$  receives a utility of  $u(V_i - P_i - K)$ .

With a slight abuse of notation, we use  $N$  and  $A$  to denote the number as well as the set of potential bidders and entrants respectively. Let  $F_{\xi_1, \xi_2}, F_{\xi_2 | \xi_1}$  denote respectively the joint and conditional distributions of generic random vectors  $(\xi_1, \xi_2)$ . We use upper cases to denote random variables and lower cases to denote their realizations. We take  $N$  as given and fixed in this section and, to simplify exposition, we drop  $N$  in notations when there is no ambiguity.

**Assumption 1** *Given  $K$ , private values  $V_i$  are i.i.d. draws from the same continuous distribution  $F_{V|K}$  that has positive densities almost everywhere on the support  $[\underline{v}, \bar{v}]$ . Entry costs across auctions are i.i.d. draws from the same distribution  $F_K$  with support  $[\underline{k}, \bar{k}]$ .*

This assumption states that bidders have symmetric, independent private values in the sense that  $\Pr(V_1 \leq v_1, \dots, V_N \leq v_N | K = k) = \prod_{i \in N} F_{V|k}(v_i)$  for all  $k$  and  $(v_1, \dots, v_N)$ . Each entrant  $i$  in a bidding stage follows the weakly dominant equilibrium strategy to drop out at his true value  $V_i$  if  $A \geq 2$ . When  $A = 1$  in the bidding stage, the lone entrant wins and pays the reserve price  $r$ .

Let  $A_{-i}$  denote the set (and the number) of entrants that  $i$  competes with if he enters. Let the reserve price be binding (i.e.  $r > \underline{v}$ ). If  $A_{-i} \neq \emptyset$ , define  $P_i \equiv \max_{j \in A_{-i}} \{V_j, r\}$  as  $i$ 's payment if he enters and wins while all competitors in  $A_{-i}$  follow their weakly dominant bidding strategies; if  $A_{-i} = \emptyset$ , define  $P_i \equiv r$ . Then  $i$ 's (random) profit in the weakly dominant strategy equilibrium is  $(V_i - P_i)_+ - K$ , where  $(\cdot)_+ \equiv \max\{\cdot, 0\}$ . Let  $\omega(k; \boldsymbol{\lambda}_{-i})$  denote expected utility for bidder  $i$  conditional on paying entry cost  $k$  and potential competitors entering with probabilities  $\boldsymbol{\lambda}_{-i} = (\lambda_j)_{j \in N \setminus \{i\}} \equiv (\lambda_1, \dots, \lambda_{i-1}, \lambda_{i+1}, \dots, \lambda_N)$ . Under Assumption 1,

$$\omega(k; \boldsymbol{\lambda}_{-i}) \equiv u(-k)F_{V|k}(r) + \int_r^{\bar{v}} h(v, k; \boldsymbol{\lambda}_{-i}) dF_{V|k}(v), \quad (1)$$

where for all  $v > r$ ,

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

with  $F_{P_i}(\cdot|k, \boldsymbol{\lambda}_{-i})$  being the distribution of  $P_i$  when  $K = k$ , and  $i$ 's potential competitors enter with probabilities  $\boldsymbol{\lambda}_{-i}$ . Due to symmetry in private value distributions across bidders,  $F_{P_i}(\cdot|k, \boldsymbol{\lambda}_{-i})$  does not change with the bidder identity  $i$ , and therefore  $\omega$  does not depend on the bidder identity  $i$ . The following lemma characterizes the symmetric mixed strategy equilibrium in the entry stage:



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

Whenever  $\omega(k; (0, \cdot, 0)) \leq u(0)$  (respectively,  $\omega(k; (1, \cdot, 1)) \geq u(0)$ ), the equilibrium entry probabilities are degenerate at 0 (respectively, 1). Thus the condition that  $\omega(k; (1, \cdot, 1)) < u(0) < \omega(k; (0, \cdot, 0))$  can be tested, as long as entry decisions are reported in data. The equilibrium entry process is non-selective, as potential bidders' entry decisions are not based on any information correlated with private values to be drawn in the bidding stage.

It is worth noting that we can generalize the model to incorporate auction heterogeneities  $Z$  that are known to all potential bidders *and* reported in data. To do this, we need to condition private value distributions on  $Z$  as well. The characterization of entry and bidding strategies in equilibrium can be derived using arguments similar to Lemma 1, for any given  $z$ . In fact, auction heterogeneities reported in data provide additional identifying power for inferring bidders' risk attitudes when entry costs are not reported in data. (See Section 5.)

We assume researchers know the number of potential bidders in auctions. This means either the data reports the number of potential bidders directly, or researchers can construct a measure of  $N$  from the data based on institutional details in the specific environments. Such an assumption holds in a variety of applications considered in the literature. For example, Li and Zheng (2009) considers highway mowing auctions by Texas Department of Transportation. In that case, the number of potential bidders for each auction is reported as the observed number of contractors who have requested the official bidding proposal for the project. Similarly, in Krasnokutskaya and Seim (2011), potential bidders in highway procurement auctions in California are defined as companies that have purchased the detailed project plan from California Department of Transportation. Such purchases are recorded for all projects in data. Athey, Levin and Seira (2011) studies US Forest Service (USFS) timber auctions. They measure the number of loggers potentially interested in an auction by the number of logging companies that had entered auctions in the same geographic area in the prior year. Such a proxy is constructed by combining data from timber auctions with the US Census data.

## 4 Inference of Risk Attitudes Using Entry Costs

We now introduce a test for bidders' risk attitudes when the data allows researchers to construct a consistent estimator for the expectation of entry costs in the data-generating process. In some contexts, a fixed amount of entry costs, sometimes also known as admission fees, is charged to all entrants and recorded in data. Such admission fees occur in contexts such as: auctions of used or new cars in the U.S. and U.K.; auctions of vintage wines in the

U.S.; and auctions of art items, etc. In other applications, entry costs are not admission fees but nevertheless are measurable, at least up to random noises through additional surveys.<sup>3</sup> That said, it is also likely that these noisy measures may systematically underestimate true costs, and taking the average of these noisy measures would lead to a biased estimator for the actual mean. Thus the test based on the sign of risk-premia would be biased towards the direction of risk-averse alternatives. We provide some simulation exercises about the impact of such bias in Section 4.4 below.

In addition to its empirical relevance, there is also a theoretical motivation for this assumption: without variations in potential competition or information about entry costs, bidders' risk attitudes can not be inferred from entry decisions and transaction prices alone without imposing additional parametric assumptions on the structure. To fix ideas, consider a simplified model where all auctions in data share the same fixed entry cost  $k$  that is not reported in data. The private value distribution is identified from the distributions of transaction prices and the number of entrants using standard arguments using order statistics; and the equilibrium entry probability is directly recovered from entry decisions. Nonetheless,  $u(\cdot)$  and the unknown entry cost  $k$  cannot be jointly identified in this case.

To see this, suppose bidders are risk-neutral with a continuous utility function  $u(\cdot)$ , entry cost  $k$ , and entry probability  $\lambda^* \in (0, 1)$ . Fix a continuous private value distribution and let it be independent from entry costs. Now consider a slightly concave utility  $\tilde{u}(\cdot) \neq u(\cdot)$  so that  $E[\tilde{u}((V_i - P_i)_+ - k) | \lambda^*] > \tilde{u}(0)$ . Because the left-hand side is continuous and strictly monotone in  $k$ , we can increase  $k$  to  $\tilde{k} > k$  so that the indifference condition is restored for  $(\tilde{u}(\cdot), \tilde{k}) \neq (u(\cdot), k)$ . That is, the equilibrium entry probability observed,  $\lambda^*$ , can be rationalized by more than one data-degenerating processes. Thus, without variations in potential competition or auction heterogeneity, inference of risk attitudes from in English auctions with endogenous entry must utilize at least some partial knowledge of entry costs.

All in all, we acknowledge that the assumption of observable entry costs used in the current section is strong. In addition to empirical and theoretical motivations above, we also hope to motivate the test under such an assumption as a useful benchmark that shows how

---

<sup>3</sup>For example, in USFS timber auctions considered in Athey, Levin and Seira (2011), entry costs for potential bidders (i.e., millers and loggers located in nearby geographic region) consist largely of information acquisition costs. These costs are incurred while performing "cruises" over the auctioned tracts to learn the distribution of diameters and heights of trees, etc. Such private cruises are standard practices institutionalized in the industry, and their costs vary little across millers and loggers. Thus cruise costs can be treated as practically identical for all potential bidders. Besides, it is plausible that researchers can construct a consistent estimator of average costs for cruising tracts with given characteristics, because cruise costs are likely to be measured up to random errors through additional field work of data collection (such as surveys or industry interviews). As cruise costs are orthogonal to bidders' private values (or realized profitability) upon entry, one could reasonably expect there to be no incentive to systematically under- or over-state cruise costs in surveys. In this case, survey data, if collected, could well be expected to provide a consistent estimator for expected entry costs.

entry behaviors can be informative about bidders' risk premia, which in turn is useful for inferring their risk attitudes.

## 4.1 Identifying Bidders' Risk Attitudes

Assume that the number of potential bidders is fixed and known to researchers. Our test for risk attitudes in this subsection builds on the simple intuition that the certainty equivalent for *risk-averse* bidders is strictly less than ex ante expected profits from entry. To fix ideas, we first show the difference between these two quantities can be recovered from observed distributions of entry decisions and transaction prices when entry costs are fixed and reported in data. Later in this subsection, we show how to extend this approach under a more practical scenario where data does not report perfectly the realizations of entry costs, but allows researchers to construct a consistent estimator for its mean.

First, we show that bidders' *ex ante* expected profits are identified. Let  $\lambda_k^*$  denote entry probabilities in the symmetric BNE when the common entry cost is  $k$ . Let  $\pi(k)$  denote the ex ante profit for a bidder  $i$  if he enters, conditional on entry cost  $k$  and that each of his potential competitors enters with probability  $\lambda_k^*$ . That is,  $\pi(k) \equiv E[(V_i - P_i)_+ - K | K = k]$ . (Strictly speaking, the definition of  $\pi(k)$  is conditional on the event that " $\lambda_j = \lambda_k^* \forall j \neq i$ ". We suppress this from the notation to simplify notations.) Note  $\pi(\cdot)$  is independent from bidder identities due to the symmetry in private value distributions. For all  $s \leq m$ , let  $V^{(s:m)}$  denote the  $s$ -th smallest among  $m$  independent draws of private values from some parent distribution  $F_{V|k}$ . Let  $F_{V^{(s:m)}|k}$  denote the distribution of this order statistic given entry cost  $k$ .

**Proposition 1** *Suppose Assumption 1 holds and entry costs are reported in data. For any  $k$  with  $0 < \lambda_k^* < 1$ ,  $\pi(k)$  is identified from bidders' entry decisions and the distribution of transaction prices.*

Proof of Proposition 1 is included in Appendix A. The intuition builds on two simple observations: First, Assumption 1 guarantees  $F_{V|k}$  can be recovered from the distribution of transaction prices under cost  $k$ . This is because, with the number of entrants  $A$  reported in data, the distribution of prices is a distribution of the second-largest order statistics out of a known number of i.i.d. draws from  $F_{V|k}$ . Second, once  $F_{V|k}$  is recovered,  $\pi(k)$  can be calculated as a known functional of the distribution of private values. It is worth noting that identification of ex ante surplus  $E[(V_i - P_i)_+ | k]$  per se does not rely on assumed knowledge of entry costs from data. Rather, it could be recovered for an unknown entry costs  $k$ , as long as researchers are aware of which auctions in data share this particular level of entry cost  $k$ .

**Proposition 2** *Suppose Assumption 1 holds. For any  $k$  such that  $0 < \lambda_k^* < 1$ ,  $\pi(k) = 0$  if and only if bidders are risk-neutral, and  $\pi(k) > 0$  (or  $< 0$ ) if and only if bidders are risk-averse (or risk-loving).*

*Proof.* Lemma 1 showed that for any such  $k$  and in a symmetric BNE, bidders enter with probability  $\lambda_k^*$ , where  $\omega(k; \lambda_k^*) \equiv E[u((V_i - P_i)_+ - k)|k, \text{“}\lambda_j = \lambda_k^* \forall j \neq i\text{”}] = u(0)$ . Thus, zero is the certainty equivalent associated with  $u(\cdot)$  and the distribution of  $(V_i - P_i)_+$  given an entry cost  $k$  and the value distribution  $F_{V|k}$ . It then follows that  $\pi(k) > 0$  if  $u'' < 0$  (bidders are risk-averse). Likewise,  $\pi(k) = 0$  (or  $\pi(k) < 0$ ) if bidders are risk-neutral (or risk-loving).  $\square$

If the exit prices by all losing bidders are reported in data, the distribution of other order statistics  $V^{(\tilde{m}:m)}$  with  $\tilde{m} \leq m - 2$  provides a source of over-identification of  $\pi(k)$ . This is because the one-to-one mappings between  $F_{V|k}$  and  $F_{V^{(\tilde{m}:m)}|k}$  exists for  $\tilde{m} \leq m - 2$  as well. Such over-identification should be exploited to improve efficiency in the estimation.

We now turn to the more practical scenario where data provide enough information to identify the *expectation* of entry costs, but do not report realized entry costs in each auction. Such a scenario is relevant, for example, where data does not report  $K$  but provides noisy measures of costs  $\tilde{K} = K + \epsilon$  with  $E(\epsilon) = 0$ . In such cases,  $\mu_K \equiv E(K) = E(\tilde{K})$  is directly identifiable.

**Corollary 1 (of Proposition 2.)** *Suppose Assumptions 1 hold. (a) If  $0 < \lambda_k^* < 1$  for all  $k \in [\underline{k}, \bar{k}]$ , then  $E[\pi(K)] = 0$  when bidders are risk-neutral, and  $E[\pi(K)] > 0$  (or  $< 0$ ) when bidders are risk-averse (or risk-loving); (b) 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 Corollary 1 is included in the appendix. The key idea underlying this result is that, with the sign of  $u''(\cdot)$  assumed fixed over its domain, the testable implications in Proposition 2 are preserved after integrating out entry costs under conditions for Corollary 1. Thus, with the expectation of entry costs assumed known in Corollary 1, the test only requires researchers to recover  $E[\pi(K)]$  (i.e., *ex ante* profits after  $K$  is integrated out). The latter is achieved by exploiting two facts. First, once conditioning on the number of entrants, bidders' expectation of their surplus  $(V_i - P_i)_+$  is independent from entry costs. This is because of the orthogonality condition between private values and the entry costs assumed in Part (b) of Corollary 1. Thus the expectation of  $(V_i - P_i)_+$  given the number of competing entrants  $A_{-i}$  can be recovered from the data as in Proposition 2. Second, the properties of the (binomial) distribution of the number of entrants can be used to relate the unconditional distribution of the number of competitors for  $i$  in the bidding stage to observed distributions.

Our approach here can be extended to allow for observed auction-level heterogeneities in data. To do so, we need to modify Proposition 1 and 2 and Corollary 1 by conditioning the assumptions and results therein on observable realizations of auction-level heterogeneities.

The main limitation of the approach in this subsection is of course its reliance on the existence of a consistent estimator for the mean of entry costs. This assumption does not hold when the noisy measures of entry costs in data are subject to systematic omissions or overstatement. In those cases, the average of these noisy measures yields a biased estimator for expected entry costs.<sup>4</sup> In Section 5 we propose an alternative test that does not require knowledge of entry costs and thus is not subject to this limitation. This is done by exploiting exogenous variations in potential competition and auction heterogeneities reported in data.

## 4.2 Test Statistic and Bootstrap Inference

We propose a statistic for testing bidders' risk attitudes under the conditions of Corollary 1. To reiterate, the entry costs  $K$  vary across auctions independently from  $(V_i)_{i \in N}$ , and for each auction the researcher only observes a noisy measure of the entry cost  $\tilde{K} = K + \epsilon$ , where  $\epsilon$  has zero mean and is independent from  $K$  and  $(V_i)_{i \in N}$ . To rule out uninteresting cases with degenerate entry behavior, we maintain that  $0 < \lambda_k^* < 1$  for all  $k \in [\underline{k}, \bar{k}]$  throughout. To simplify exposition, we fix  $N$  and suppress it from notations throughout the section and in Appendix B. The following conditions are maintained in our derivation of the limiting distribution of the test statistic.

**Assumption R.** (i) Across all auctions, the  $(N + 2)$ -vector  $((V_i)_{i \in N}, K, \epsilon)$  are *i.i.d.* draws from the same joint distribution. (ii) All coordinates in  $(V_i)_{i \in N}$ ,  $K$  and  $\epsilon$  are mutually independent with finite second moments and  $E(\epsilon) = 0$ . (iii) Within each auction, the private values  $V_i$  are independent draws from the same continuous marginal distribution  $F_V$ , which has positive densities over  $[\underline{v}, \bar{v})$ , and has a probability mass at  $\bar{v}$ . (iv) The reserve price  $r$  is fixed and binding ( $r > \underline{v}$ ) throughout the data.

That  $V_i$  is independent from  $K$  is required for identifying the ex ante risk premium  $E[\pi(K)]$ , as shown in Corollary 1. That  $\epsilon$  has a zero mean ensures the sample mean of  $\tilde{K}$  in data is an unbiased and consistent estimator for  $E(K)$ . Finite second moments of  $(K, \epsilon)$  are needed for applying a multivariate central limit theorem while deriving the limiting distribution of our test statistic. That  $F_V$  has a probability mass at the upper bound  $\bar{v}$  is a technical condition that simplifies the proof for weak convergence of our estimator for  $F_V$  below. It is not needed in our argument above for recovering bidders' risk attitudes. In what follows, we let  $\int_r^{\bar{v}} \zeta(s) ds$  be a shorthand for the improper integral  $\lim_{v \uparrow \bar{v}} \int_r^v \zeta(s) ds$ , where  $\zeta$  is a generic integrable function, for the rest of this section and in Appendix B. While correlation between noises  $\epsilon$  and entry costs  $K$  is allowed in Corollary 1, the orthogonality in

---

<sup>4</sup>If the direction of bias is known to be upward then averaging these noisy measures gives a consistent estimator for some upper bound on expected entry costs. One can still test the null that *ex ante* surplus  $E[(V_i - P_i)_+ | k]$  does not exceed the upper bound of expected entry costs. If the null is rejected then we can conclude bidders are risk averse. Nonetheless such a test has limited value in practice, for it is inconclusive about risk attitudes when the null is not rejected.

Assumption R-(ii) helps simplify derivations of the asymptotic property of the test statistic. That the reserve price is binding means  $F_V(r) > 0$ , which is empirically relevant in lots of contexts.

Our goal is to infer which of the three hypotheses below is best supported by data:

$$\begin{aligned} H_0 &: \tau_0 = 0 \text{ (risk-neutral);} \\ H_A &: \tau_0 > 0 \text{ (risk-averse); and } H_L : \tau_0 < 0 \text{ (risk-loving).} \end{aligned}$$

where  $\tau_0 \equiv E[\pi(K)]$ . The data contain  $T$  independent auctions, each of which is indexed by  $t$  and involves  $N$  potential bidders. Let  $A_t$  denote the number of entrants in auction  $t$ . Let  $W_t$  define the transaction price in auction  $t$ . If the object is sold, then  $W_t = \max\{r, V^{(A_t-1:A_t)}\}$  when  $A_t \geq 2$  and  $W_t = r$  when  $A_t = 1$ . If the object is not sold, then define  $W_t < r$ .

Our test statistic  $\hat{\tau}_T$  is a multi-step estimator for the unconditional expectation of a bidder's profit  $E[\pi(K)]$  based on the sample analog principle. First, for  $m \geq 2$ , estimate the distribution of transaction prices:

$$\hat{F}_{W|m,T}(s) \equiv \frac{1}{T} \sum_{t \leq T} 1\{W_t \leq s, A_t = m\} / \frac{1}{T} \sum_{t \leq T} 1\{A_t = m\}$$

for any  $r \leq s < \bar{v}$ . Then for any  $m \geq 2$ , estimate the private value distribution by:

$$\hat{F}_{V,T}(s) \equiv \frac{1}{N-1} \sum_{m=2}^N \phi_m^{-1}(\hat{F}_{W|m,T}(s)) \text{ for } s \geq r, \quad (3)$$

where  $\phi_m(t) \equiv t^m + mt^{m-1}(1-t)$ . Next, define  $\varphi_a \equiv E[(V_i - P_i)_+ | A_{-i} = a]$  for  $0 \leq a \leq N-1$ . Estimate  $\varphi_a$  using the sample analog principle:<sup>5</sup>

$$\hat{\varphi}_{T,a} \equiv \int_r^{\bar{v}} \left[ \hat{F}_{V,T}(s) \right]^a \left[ 1 - \hat{F}_{V,T}(s) \right] ds, \quad (4)$$

where the integral can be calculated using mid-point approximations. Estimate the distribution of  $A_{-i}$  by the column vector  $\hat{\rho}_T \equiv [\hat{\rho}_{T,0}, \hat{\rho}_{T,1}, \dots, \hat{\rho}_{T,N-1}]'$  where:

$$\hat{\rho}_{T,a} \equiv \frac{1}{T} \sum_{t \leq T} \left[ \frac{N-a}{N} 1\{A_t = a\} + \frac{a+1}{N} 1\{A_t = a+1\} \right]$$

for  $0 \leq a \leq N-1$ . (See equation (A5) in Appendix A for details.) Finally, calculate the test statistic by:

$$\hat{\tau}_T \equiv \sum_{a=0}^{N-1} \hat{\varphi}_{T,a} \hat{\rho}_{T,a} - \frac{1}{T} \sum_{t \leq T} \tilde{K}_t.$$

We now derive the limiting distribution of  $\sqrt{G}(\hat{\tau}_T - \tau_0)$ . Let “ $\rightsquigarrow$ ” denote weak convergence of stochastic processes in a normed space. Let  $F_{\bar{V}}$  denote  $\{F_V(t) : t \in [r, \bar{v}]\}$  and  $\hat{F}_{\bar{V},T}$  denote  $\{\hat{F}_{V,T}(t) : t \in [r, \bar{v}]\}$ . Denote the limiting distribution of  $\sqrt{T}(\hat{\rho}_T - \rho)'$  and  $\sqrt{T}[\frac{1}{T} \sum_{t \leq T} \tilde{K}_t - \mu_K]$  by  $\mathcal{N}'_{\rho}$  and  $\mathcal{N}_{\mu}$  respectively. We characterize the covariance between

<sup>5</sup>See Appendix B for a proof that, under Assumption R,  $\varphi(a) = \int_r^{\bar{v}} [F_V(s)]^a [1 - F_V(s)] ds$  for each  $a \geq 0$ .

$\mathcal{N}'_\rho$  and  $\mathcal{N}_\mu$  in Lemma B5 in Appendix B. In Lemma B6 of Appendix B, we show under the mild conditions in Assumption R,  $\sqrt{T} \left( \hat{F}_{\hat{V},T} - F_{\hat{V}} \right) \rightsquigarrow \mathbb{G}_V$  where  $\mathbb{G}_V$  is a zero-mean Gaussian process, and characterize covariance kernel of  $\mathbb{G}_V$  as well as its covariance with  $[\mathcal{N}'_\rho, \mathcal{N}_\mu]$ .

Let  $\mathcal{S}_{[r,\bar{v}]}$  denote the set of positive, bounded and integrable Cadlag functions over  $[r, \bar{v}]$ .<sup>6</sup> For any  $F \in \mathcal{S}_{[r,\bar{v}]}$ , define  $\varphi(F) \equiv (\varphi_a(F))_{a=0}^{N-1}$ , where  $\varphi_a(F) \equiv \int_r^{\bar{v}} [F(s)^a - F(s)^{a+1}] ds$  for all  $a \geq 0$ . By definition,  $\varphi(\hat{F}_{\hat{V},T}) = (\hat{\varphi}_{T,a})_{a=0}^{N-1} \equiv \hat{\varphi}_T$ . It follows from van der Vaart (1998) (Lemma 20.10 on page 298) that  $\varphi$  is Hadamard differentiable at  $F_{\hat{V}}$ , tangentially to  $\mathcal{S}_{[r,\bar{v}]}$ . The Hadamard derivative is given by :

$$D_{\varphi, F_{\hat{V}}}(h)(a) \equiv \int_r^{\bar{v}} [aF_{\hat{V}}(s)^{a-1} - (a+1)F_{\hat{V}}(s)^a] h(s) ds \text{ for } 0 \leq a \leq N-1$$

for any  $h \in \mathcal{S}_{[r,\bar{v}]}$ .

**Proposition 3** *Suppose Assumption R holds with  $0 < \lambda_k^* < 1$  for all  $k \in [\underline{k}, \bar{k}]$ . Then*

$$\sqrt{T}(\hat{\tau}_T - \tau_0) \rightsquigarrow \mathcal{N}_\tau$$

where  $\mathcal{N}_\tau \equiv \rho' D_{\varphi, F_{\hat{V}}}(\mathbb{G}_V) + \varphi' \mathcal{N}_\rho - \mathcal{N}_\mu$  is a univariate normal distribution with zero mean.

For a given level  $\alpha$ , let  $\hat{c}_{1-\alpha/2, T}$  denote an estimator for the  $100 \cdot (1 - \alpha/2)$ -percentile of the limiting distribution of  $\sqrt{T}(\hat{\tau}_T - \tau_0)$  using bootstrap procedures. (See Appendix C for the definition of  $\hat{c}_{1-\alpha/2, T}$ .) The decision rule for the test is to reject  $H_0$  in favor of  $H_A$  (or  $H_L$ ) if  $\sqrt{T}\hat{\tau}_T > \hat{c}_{1-\alpha/2, T}$  (or if  $\sqrt{T}\hat{\tau}_T < -\hat{c}_{1-\alpha/2, T}$ ); and do not reject  $H_0$  if  $-\hat{c}_{1-\alpha/2, T} \leq \sqrt{T}\hat{\tau}_T \leq \hat{c}_{1-\alpha/2, T}$ .

The next proposition shows the test is consistent against fixed alternatives, and attains the correct asymptotic level. Let  $\Pr(\hat{\tau}_T \leq \cdot | \tau_0 = \tau_0^*)$  denote the distribution of  $\hat{\tau}_T$  conditioning on the true value for  $\tau_0$  in data-generating process being  $\tau_0^*$ .

**Proposition 4** *Suppose Assumption R holds with  $0 < \lambda_k^* < 1$  for all  $k \in [\underline{k}, \bar{k}]$ . Then*

$$\lim_{T \rightarrow +\infty} \Pr \left( \sqrt{T}\hat{\tau}_T > \hat{c}_{1-\alpha/2, T} | \tau_0 = \tau_0^* \right) = 1 \quad \forall \tau_0^* > 0; \quad (5)$$

$$\lim_{T \rightarrow +\infty} \Pr \left( \sqrt{T}\hat{\tau}_T < -\hat{c}_{1-\alpha/2, T} | \tau_0 = \tau_0^* \right) = 1 \quad \forall \tau_0^* < 0; \quad (6)$$

$$\lim_{T \rightarrow +\infty} \Pr \left( \sqrt{T}\hat{\tau}_T > \hat{c}_{1-\alpha/2, T} \text{ or } \sqrt{T}\hat{\tau}_T < -\hat{c}_{1-\alpha/2, T} | \tau_0 = 0 \right) = \alpha. \quad (7)$$

Two facts are used for showing these results. First, the empirical distribution of  $\sqrt{T}(\hat{\tau}_{T,b} - \hat{\tau}_T)$  provides a consistent estimator for the finite sample distribution of  $\sqrt{T}(\hat{\tau}_T - \tau_0)$  under the stated conditions, which is verified in Appendix C by using results from Beran and Ducharme (1991). Second,  $\sqrt{T}\tau_0$  is zero under the null but diverges to positive (or negative) infinity under the alternative.<sup>7</sup>

<sup>6</sup>A Cadlag function is right-continuous with left limits.

<sup>7</sup>Our bootstrap inference uses an asymptotically non-pivotal statistic  $\sqrt{T}(\hat{\tau}_T - \tau_0)$ . One could construct

### 4.3 Monte Carlo Simulations

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 set 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. realized private 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  $\epsilon$  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  $\gamma > 1$ ). We experiment with the number of potential bidders  $N = 4$  or  $N = 5$ , and sample sizes  $T = 1, 500$  or  $T = 3, 000$ .

To improve the test performance in small samples, we modify our estimator for risk-premium  $\hat{\tau}_T$  slightly by replacing the averaged estimators for the value distribution  $F_V$  in (3) with inverse-variance-weighted average estimators. That is, while constructing  $\hat{\tau}_T$ , we replace  $\hat{F}_{V,T}(s)$  in (4) by  $\tilde{F}_{V,T}(s) \equiv \frac{1}{N-1} \sum_{m=2}^N \hat{\beta}_m(s) \phi_m^{-1}(\hat{F}_{W|m,T}(s))$  for  $s \geq r$ , with  $\hat{\beta}_m(s) \equiv (\hat{\sigma}_{m,T}^2(s))^{-1} / \sum_{m'=2}^N (\hat{\sigma}_{m',T}^2(s))^{-1}$  and  $\hat{\sigma}_{m,T}^2$  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 (as explained in Section 4.2) in our inference procedures below.

[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  $\gamma$  between 0.75 and 1.25 (with grid width 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  $\gamma$  and sample size  $T = 1, 500$  or  $3, 000$ , we simulate  $S = 250$  samples.

---

asymptotically pivotal statistics using the pre-pivoting approach. This would help attain asymptotic refinements in the approximation of test statistic distribution relative to first-order asymptotic approximation or bootstrap using asymptotically non-pivotal statistics. This is computationally intensive due to bootstrap iterations and therefore we do not pursue this approach here.



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.<sup>8</sup> 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  $(N, \gamma)$  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  $\alpha$  under the null, and reasonably high probabilities for rejecting the null in favor of the correct alternative when  $\gamma \neq 1$ . When  $\gamma \neq 1$ , the probability for “Type-III” error (i.e. rejecting the null in favor of a wrong alternative) is zero across all specifications and sample sizes.

---

<sup>8</sup>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.

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  $\gamma$  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  $\alpha$ , 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 alternative seems unambiguous: For all  $\gamma \neq 1$  and given any sample size, the power appears to be higher with  $N = 4$  for all  $\alpha$ . 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 above pattern.

#### 4.4 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 underestimates 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 rejection 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).

## 4.5 Mismeasurement of $N$

The number of potential bidders  $N$  could also be mismeasured in data for several distinct reasons, each of which has a different impact on the performance of our test.

First, under-measurement of  $N$  could happen due to finite sample limitations. For example, consider the case where equilibrium entry probabilities are low and the set of potential bidders is measured as the union of entrants observed in the sample. In such a case, a potential bidder may not be counted in because it never entered as an active bidder in the finite sample. Our estimator for entry probabilities is thus biased upward under this mismeasurement, because potential bidders who decide not to enter in any auction reported in data are not counted in the denominator. On the other hand, this mismeasurement does not impact the estimation of the private value distribution, since the number of entrants is correctly reported in data.

To see how this affects our test, recall bidders' risk premia are weighted sums of interim profits from entry. Such an under-measurement in  $N$  has no impact on the estimation of these interim profits, which only depend on the private value distribution, the number of competing entrants  $A_{-i}$ , and the expectation of entry costs. On the other hand, it does affect estimates for the weights, or probability masses for  $A_{-i}$ , through its impact on estimates of entry probabilities. Unfortunately the sign of the resultant bias in the estimator for risk-premia is indeterminate, for it depends on the way those interim profits vary with  $A_{-i}$ , and how the distribution of  $A_{-i}$  changes with entry probabilities.<sup>9</sup> Nevertheless, it is worth

---

<sup>9</sup>That the estimators for entry probabilities are known to be biased upward under this mismeasurement is not sufficient for deriving the sign of bias in our risk-premium estimator. This sign is determined by

noting that, as the real  $N$  increases in DGP, the effect of under-measurement on the bias in the estimator for entry probabilities and risk premium decreases.

The second possibility for mismeasurement is over-counting  $N$  due to problematic assumptions. Again, consider the case where potential bidders are measured as the union of entrants observed in data. An implicit assumption is that a bidder who shows up in one auction is by default also a potential bidder in other auctions in data. This risks over-counting  $N$  when there are bidders who pay the costs to enter in some auctions but are not interested as a potential bidder in the others, say, due to budget or time constraints. Such over-counting leads to a *downward* bias in the estimator for entry probabilities but has no bearing on the estimation of interim profits. Again, the ultimate impact of such a bias on test performance is indeterminate, but diminishes as the real  $N$  increases.

Yet a third scenario is that the number of potential bidders reported in data, denoted by  $\tilde{N}$ , is a persistent random under-measurement of the real  $N$  in the DGP. In this case, the conditional distribution of  $\tilde{N}$  given  $N$ , denoted by  $\tilde{g}(\tilde{N}|N)$ , is an additional parameter to be identified. This differs qualitatively from the first possibility (under-measurement due to finite sample limitations) in that the distribution  $\tilde{g}$  is a model primitive that does not vary with the sample size. An, Hu and Shum (2010) studied the estimation of a related model of first-price auctions where the number of actual bidders are misclassified due to the truncation by binding reserve prices. It turns out we can apply their argument to the current context to identify the joint distribution of  $(N, \tilde{N})$ , and the distribution of transaction prices given  $N$ . Consequently, the entry probabilities conditional on  $N$  are also identified and tests for risk attitudes can be constructed as before.

For easy exposition of how the argument in An, Hu and Shum (2010) can be applied here, fix entry costs and let  $\tilde{N} = \tilde{f}(N, \tilde{\nu})$  and  $N = f(Z, \nu)$ , where  $Z$  are some instruments that affect bidders' potential participation but not the distributions of their private values. Assume bidders' private values,  $\tilde{\nu}$  and  $Z$  are mutually independent conditional on  $N$ . This implies that the transaction price  $W$  is independent from  $\tilde{N}$  and  $Z$  given  $N$ , and that  $\tilde{N}$  is independent from  $Z$  given  $N$ . Provided the joint support of  $(Z, \tilde{N})$  satisfy a mild full-rank condition, the conditional distribution of  $\tilde{N}$  given  $N$ , the conditional distribution of  $W$  given  $N$ , and the joint distribution of  $(Z, N)$  are jointly identified using a typical matrix diagonalization argument. (See Theorem 1 in An, Hu and Shum (2010) for a proof.) Knowledge of these distributions, together with the directly identifiable distribution of  $W$  given the number of entrants, implies that the entry probabilities are identified and consistently estimable. The proposed test for bidders' risk attitudes can then be constructed. We leave the implementation of a test based on such an idea for future research.

---

primitive and endogenous model elements (i.e. the utility function, the private value distribution, and the true equilibrium entry probabilities).

## 5 Inference of Risk Attitudes with Unobserved Entry Costs

In lots of empirical environments, auctions reported in data differ in observed characteristics of the auctioned object. Besides, the number of potential bidders  $N$  may vary across auctions. In this section, we show how to use these sources of variations to infer bidders' risk attitudes, when entry costs known to bidders are *not* reported in data.

**Assumption 2**  $V_i = h(Z) + \eta_i$  for all  $i$ , where  $Z$  is a vector of auction characteristics reported in data, and  $\eta_i$  are *i.i.d.* draws from some distribution  $F_\eta$  that is independent from  $Z$  and  $N$ .

The function  $h(\cdot)$  is a model primitive unknown to the econometrician. It is directly identifiable under a location normalization  $\mathbb{E}[\eta] = 0$ . To see this, let  $F_{V|Z}$  denote the distribution of  $V_i$  conditional on  $Z$ , where the subscript  $i$  is dropped due to symmetry. By the same argument used in Proposition 1 in Section 4.1,  $F_{V|Z}$  is identified from the distribution of transaction prices  $F_{W|Z}$ . Assumption 2 and  $\mathbb{E}[\eta] = 0$  then imply  $h(z)$  is identified as  $h(z) = \mathbb{E}[V_i|z] = \int v dF_{V|z}(v)$  for all  $z$ .

Independence of  $\eta$  from  $N$  in Assumption 2 is analogous to the assumption of exogenous variations in  $N$  in Haile, Hong and Shum (2004) and Guerre, Perrigne and Vuong (2009). As shown in Guerre, Perrigne and Vuong (2009), it is possible to identify bidders' risk attitudes even when  $N$  is endogenous, as long as the data contains valid instruments once conditioning on some control variables. We provide detailed discussions about this in Section 5.4.1.

Under Assumption 2, a test for risk attitudes can be constructed even when entry costs are not reported in data. The idea is to exploit the fact that bidders' risk attitudes affect how entry probabilities vary with  $Z$  and  $N$  in equilibrium. To illustrate this idea, it is instructive to investigate the indifference condition in equilibrium. Such conditions equate bidders' utility from the certainty equivalent,  $u(0)$ , with their ex ante utility from entry. By construction ex ante utility from entry is a weighted average of "*interim*" utility, which conditions on the number of competing entrants with expectations taken with respect to private values. Entry probabilities enter ex ante utilities in the indifference condition through the weights assigned to interim utilities. These weights correspond to probability mass functions for the number of competing entrants. As  $Z$  and  $N$  vary, the entry probabilities (and therefore the weights) change endogenously in order to respect indifference condition in equilibrium.

The idea of our test builds on the following simple observation. If bidders are risk-neutral, then variations in  $Z$  induce the same rate of changes in interim utilities as that in expected private values. On the other hand, this equality fails when bidders are risk-averse (or risk-loving) due to the decreasing (or increasing) rate of increase in utility. Under the

orthogonality and additive separability conditions in Assumption 2, the rate of changes in expected private values is over-identified as the ratio between changes in expected transaction prices.<sup>10</sup> Thus, a test for bidders' risk attitudes would be feasible if the rate of changes in interim utilities can be related to the distribution of entry decisions and transaction prices.

We construct a linear system that characterizes the indifference conditions for various  $Z$  and  $N$ . Under Assumption 2, the number of unknowns in the system (i.e. the interim utilities) increases at the same pace as the number of equations. As long as the matrix of coefficients in the linear system (i.e. weights for interim utilities under various  $Z$  and  $N$ ) has full-rank, we can recover the rates of changes in interim utilities from the distribution of prices and entry decisions.<sup>11</sup> This allows us to conduct a test for risk attitudes based on the afore-mentioned idea.

## 5.1 Identification

Suppose the entry cost  $k$  is fixed across auctions but not reported in data.<sup>12</sup> For easy exposition, assume there is no binding reserve price. Let  $\lambda_{z,n}$  denote bidders' equilibrium entry probabilities in auctions with  $Z = z$  and the number of potential bidders  $N = n$ . For  $0 \leq a \leq n - 1$ , define:

$$\rho_{a,z,n} \equiv \Pr(A_{-i} = a | Z = z, N = n) = \binom{n-1}{a} (\lambda_{z,n})^a (1 - \lambda_{z,n})^{n-a-1}. \quad (8)$$

In the mixed-strategy equilibrium in the entry stage,

$$u(0) = E[u((V_i - P_i)_+ - k) | Z = z, N = n] = \sum_{a=0}^{n-1} \psi_a(z) \rho_{a,z,n} \quad (9)$$

where  $\psi_a(z) \equiv E[u((V_i - P_i)_+ - k) | A_{-i} = a, Z = z]$ , and  $A_{-i}, V_i, P_i$  are defined as in Section 3. Under Assumptions 2,  $(V_i - P_i)_+$  does not depend on  $n$  given  $z$  and the number of competing entrants  $A_{-i} = a$ . For all  $a \geq 1$ ,  $\psi_a(z)$  is defined as:

$$\psi_a(z) \equiv E[u((V_i - P_i)_+ - k) | Z = z, A_{-i} = a] = E\left[u\left(\left(\eta_i - \eta^{(a:a)}\right)_+ - k\right)\right],$$

where  $\eta^{(a:a)}$  is the largest among  $a$  independent draws from  $F_\eta$ . Also, by construction,

$$\psi_0(z) = E[u(V_i - k) | Z = z] = E[u(h(z) + \eta_i - k)]$$

---

<sup>10</sup> Alternatively, we could recover the rate of changes in  $E(V_i | Z)$  due to  $Z$  from the conditional distribution  $F_{V_i | Z}$ , which is over-identified from the distribution of transaction prices conditional on  $Z$  and the number of entrants.

<sup>11</sup> See Assumption 3 and discussions after Lemma 2 below for details about the sufficient rank condition.

<sup>12</sup> We could generalize this section by allowing entry costs to vary across auctions in the data-generating process as well. To do this, we could let the vector of auction heterogeneities consist of two subvectors  $(Z, \tilde{Z})$ , and let  $k$  be an *unknown* function of  $\tilde{Z}$  alone. Our inference method below applies as long as we could condition on  $\tilde{Z}$  and exploit the variations in  $Z$ .



where the second equality follows from independence between  $\eta_i$  and  $z$ . Clearly  $\psi_a(z)$  does not depend on  $z$  for  $a \geq 1$ ; thus we let  $\psi_a^* \equiv \psi_a(z)$  for all  $a \geq 1$  and any  $z$ .

**Assumption 3** *There exists  $z, z', z''$  and  $n$  on the support of  $(Z, N)$  in the data-generating process such that  $0 < \lambda_{z,s} < 1$  for all  $2 \leq s \leq n$ , and  $\lambda_{z',n}, \lambda_{z'',n} \in (0, 1)$ .*

Assumption 3 requires there be enough variation in auction characteristics and in potential competition in data. The condition that entry probabilities are in the interior of  $(0, 1)$  rules out uninteresting cases with degenerate entry behaviors. This condition can be directly verified from data in principle.

The characterization of entry probabilities for  $Z = z$  and  $N = s$  for  $2 \leq s \leq n$  leads to a system of  $(n - 1)$  equations:

$$\psi_0(z)\rho_{0,z,s} + \sum_{a=1}^{s-1} \psi_a^* \rho_{a,z,s} = u(0) \text{ for } s = 2, \dots, n. \quad (10)$$

We also have additional equations from auctions with different observed features  $z'$  or  $z''$  and  $N = n$ . That is,

$$\psi_0(z')\rho_{0,z',n} + \sum_{a=1}^{n-1} \psi_a^* \rho_{a,z',n} = u(0); \text{ and likewise with } z' \text{ replaced by } z''. \quad (11)$$

Stacking the system of equations from (10)-(11) and moving the terms that involve  $\psi_1^*$  to the right, we have:

$$\underbrace{\begin{pmatrix} \rho_{0,z'',n} & 0 & 0 & \rho_{2,z'',n} & \cdots & \rho_{n-1,z'',n} \\ 0 & \rho_{0,z',n} & 0 & \rho_{2,z',n} & \cdots & \rho_{n-1,z',n} \\ 0 & 0 & \rho_{0,z,2} & 0 & \cdots & 0 \\ 0 & 0 & \rho_{0,z,3} & \rho_{2,z,3} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \rho_{0,z,n} & \rho_{2,z,n} & \cdots & \rho_{n-1,z,n} \end{pmatrix}}_{(n+1)\text{-by-}(n+1)} \underbrace{\begin{pmatrix} \psi_0(z'') \\ \psi_0(z') \\ \psi_0(z) \\ \psi_2^* \\ \vdots \\ \psi_{n-1}^* \end{pmatrix}}_{(n+1)\text{-by-}1} = \underbrace{\begin{pmatrix} u(0) \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ u(0) \end{pmatrix}}_{(n+1)\text{-by-}1} - \underbrace{\begin{pmatrix} \rho_{1,z'',n} \\ \rho_{1,z',n} \\ \rho_{1,z,2} \\ \vdots \\ \vdots \\ \rho_{1,z,n} \end{pmatrix}}_{(n+1)\text{-by-}1} \psi_1^*. \quad (12)$$

The linear system has  $n + 2$  unknowns (i.e.  $\psi_0(z'')$ ,  $\psi_0(z')$ ,  $\psi_0(z)$  and  $\{\psi_a^* : 1 \leq a \leq n - 1\}$ ) and  $n + 1$  equations. Nonetheless, the next lemma shows how (12) can be used to relate the ratio of changes in interim utilities  $\psi_0$  to entry probabilities as  $z$  varies.

Under Assumption 3, the coefficient matrix in (12) has a full rank of  $n + 1$ . (See the proof of Lemma 2 in Appendix A.) Now replace  $u(0)$  by 0 and replace  $\psi_1^*$  by some arbitrary constant  $c$ , and solve for the  $n + 1$  unknowns  $(\psi_0(z'')$ ,  $\psi_0(z')$ ,  $\psi_0(z)$ ,  $\{\psi_a^* : 2 \leq a \leq n - 1\})$  in (12). Denote the unique solutions by  $(\tilde{\psi}_{0,z''}, \tilde{\psi}_{0,z'}, \tilde{\psi}_{0,z}, \{\tilde{\psi}_a : 2 \leq a \leq n - 1\})$ .

**Lemma 2** *Suppose Assumptions 2-3 hold. If  $h(z') \neq h(z'')$ , then*

$$\frac{\psi_0(z') - \psi_0(z)}{\psi_0(z'') - \psi_0(z')} = \frac{\tilde{\psi}_{0,z'} - \tilde{\psi}_{0,z}}{\tilde{\psi}_{0,z''} - \tilde{\psi}_{0,z'}}. \quad (13)$$

Assumption 3 is sufficient but not necessary for (13) to hold under Assumption 2. We could relax Assumption 3 and only require the data-generating process to contain enough variations in  $Z$  and  $N$  so that a linear system similar to (12) can be constructed for at least three different values of  $Z$  with the coefficient matrix in (12) being non-singular.<sup>13</sup> The next proposition states a testable implication of bidders' risk attitudes when entry costs are not recorded in data.

**Proposition 5** (a) *Under Assumption 2,*

$$h(z_1) - h(z_2) = E(W|Z = z_1, A = a) - E(W|Z = z_2, A = a)$$

for any  $(z_1, z_2)$  and  $a \geq 2$ . (b) *Suppose Assumptions 2-3 hold for  $(z, z', z'')$  such that  $h(z'') > h(z') > h(z)$ . Then*

$$\frac{\tilde{\psi}_{0,z'} - \tilde{\psi}_{0,z}}{\tilde{\psi}_{0,z''} - \tilde{\psi}_{0,z'}} \left\{ \begin{array}{l} > \\ = \\ < \end{array} \right\} \frac{h(z') - h(z)}{h(z'') - h(z')} \text{ iff } u'' \left\{ \begin{array}{l} < \\ = \\ > \end{array} \right\} 0. \quad (14)$$

To construct a test based on Proposition 5, one needs to locate a triple  $(z, z', z'')$  such that with  $h(z'') > h(z') > h(z)$ . Under Assumption 2, this sequence of strict inequalities is equivalent to  $E[W|z'', A = a] > E[W|z', A = a] > E[W|z, A = a]$  for all  $a \geq 2$ . (See the proof of part (a) of Proposition 5.) Thus such a triple can be found using observed distribution of prices conditional on  $Z$  and the number of entrants. Note this does not require the location normalization  $E[\eta_i] = 0$ .

In some other contexts, it is possible to locate such a triple using the shape restrictions on  $h$  that are known *a priori* to researchers. For example, economic theory or common sense sometimes may restrict the function  $h(\cdot)$  to be monotonic in one of the coordinates in  $Z$ , or may suggest  $F_{V|Z}$  is stochastically ordered among a known triple  $(z, z', z'')$ . In such cases, the choices of the triple to be used in the test are immediate. In the following subsection, we assume a triple  $(z, z', z'')$  with  $h(z'') > h(z') > h(z)$  is known and fixed.<sup>14</sup>

---

<sup>13</sup>Specifically, a weaker sufficient condition for Lemma 2 is as follows: "There exists  $J \geq 3$  values of  $Z$ , denoted  $\{z^j : 1 \leq j \leq J\}$ , and  $J$  overlapping sets of integers on the support of  $N$ , each of which is denoted  $\zeta^j \equiv \{n^{j,1}, n^{j,2}, \dots, n^{j,M_j}\}$  with  $M_j \equiv \#\{\zeta^j\}$ , such that the matrix of coefficients for  $\{\psi_0(z^j) : j \leq J\}$  and  $(\psi_2^*, \dots, \psi_{\bar{n}-1}^*)$  (where  $\bar{n} \equiv \max\{\cup_{j \leq J} \{\zeta^j\}\}$ ) has full-rank in the linear system of  $\sum_{j \leq J} M_j$  equations characterizing the equilibrium entry probabilities for various  $Z$  and  $N$ ." This condition necessarily requires  $\sum_{j \leq J} M_j \geq \bar{n} + J - 2$ . Note Assumption 3 is a special case of this condition with  $J = 3$ ,  $(z^1, z^2, z^3) = (z, z', z'')$ ,  $\zeta^1 = \{2, 3, \dots, n\}$ , and  $\zeta^2 = \zeta^3 = \{n\}$ .

<sup>14</sup>In this case, we need a preliminary step for selecting a triple satisfying the inequalities, based on comparing the estimates of  $E(W|Z, a)$ . We leave issues such as how to account for the impact of pretesting in inference to future research.

## 5.2 Test Statistic

We now construct a test statistic based on Proposition 5 and some triple  $(z, z', z'')$  and some  $n$  known to satisfy Assumption 3. First, estimate the coefficient matrix on the right-hand side and the coefficient vector on the left-hand side in (12) by plugging in sample analogs of the entry probabilities into (8). Next, replace  $u(0)$  with 0 and  $\psi_1^*$  with an arbitrary negative constant  $c$  in (12) and solve it to get estimates for  $(\tilde{\psi}_{0,z''}, \tilde{\psi}_{0,z'}, \tilde{\psi}_{0,z}, \{\tilde{\psi}_a : 2 \leq a \leq n-1\})$ .<sup>15</sup> Denote these estimates by  $\hat{\psi}_0(z)$ ,  $\hat{\psi}_0(z')$ ,  $\hat{\psi}_0(z'')$  and  $\{\hat{\psi}_a^* : 2 \leq a \leq n-1\}$  respectively. Our test statistic is

$$\tilde{\tau} \equiv \frac{\hat{\psi}_0(z') - \hat{\psi}_0(z)}{\hat{\psi}_0(z'') - \hat{\psi}_0(z')} - \hat{R}$$

where

$$\hat{R} \equiv \sum_{a=2}^{n-1} \hat{q}_a \left( \frac{\hat{E}(W|z', a) - \hat{E}(W|z, a)}{\hat{E}(W|z'', a) - \hat{E}(W|z', a)} \right) \quad (15)$$

with  $\hat{E}(W|z, a)$  being kernel estimators for expected transaction prices given  $z$  and the number of entrants  $A = a$ ; and  $\{\hat{q}_a : 2 \leq a \leq n-1\}$  are data-dependent weights. In the simulations below, we use  $\hat{q}_a \equiv \hat{\sigma}_a^{-2} / (\sum_{s=2}^{n-1} \hat{\sigma}_s^{-2})$ , where  $\hat{\sigma}_a$  is the standard error for the ratio of differences in the parenthesis on the right-hand side of (15).

We now briefly discuss the asymptotic property of our test statistic  $\tilde{\tau}$ . It is instructive to look at the case where the support of  $Z$  is discrete, which is also what we consider in the simulations of Section 5.3 below. In this case,  $\tilde{\tau}$  is by construction a smooth function of sample averages. (To see this, note  $\hat{\lambda}_{z,n}$  are sample averages and  $\hat{E}(W|z, a)$  are ratios whose numerators and denominators are sample averages. It then follows that  $\hat{\psi}_0(z)$ ,  $\hat{\psi}_0(z')$ ,  $\hat{\psi}_0(z'')$  and  $\hat{R}$  are smooth functions of sample averages.) Under Assumptions 2 and 3 and the conditions on finite second moments of transaction prices, the Delta Method can be applied to show that  $\tilde{\tau}$  converges at the parametric rate to its population counterpart  $\tau_*$ , which is defined as the difference between the two ratios compared in (14). Furthermore, the asymptotic distribution of  $\tilde{\tau}$  is normal with a zero mean and a finite variance.

A test for risk attitudes follows a procedure similar to that of Section 4.2, with critical values estimated using bootstrap resampling. With  $\tilde{\tau}$  being a smooth function of sample averages, the bootstrap estimator for the exact finite sample distribution of  $\tilde{\tau}$  is consistent. For the same reason, the test using bootstrap critical values is expected to perform at least as well as one based on first-order asymptotic approximation both in terms of the errors in rejection probabilities and its asymptotic power.

---

<sup>15</sup>To identify the ratio of differences between interim utilities  $\psi_0$  under various  $z$ , the sign of the chosen constant  $c$  does not matter. However, using a negative  $c$  has the additional benefit of recovering the correct ordering of  $\psi_0$  under different  $z$ . (See the proof of Lemma 2 in Appendix A for details.)

### 5.3 Monte Carlo Simulations

This subsection presents some Monte Carlo evidence for the performance of the test above in finite samples. The DGP is as follows. The distribution of auction characteristics  $Z$  is multinomial over a discrete support  $Z \in \{1, 2, 3\}$  with equal probability masses. Upon entry, a bidder  $i$ 's private value is  $\beta_1 + \beta_2 Z + \eta_i$ , where  $\beta \equiv (\beta_1, \beta_2)$  are parameters to be experimented with and  $\eta_i$  are i.i.d draws from a uniform distribution over  $[-2, 2]$ . The support of potential bidders is  $\{2, 3, 4, 5, 6\}$ . Conditional on  $Z$ , the distribution of  $N$  is stochastically increasing:  $\Pr(N = n|Z = 1) = \frac{1}{3}$  for  $n = 2$  and  $\frac{1}{6}$  for all  $n \neq 2$ ;  $\Pr(N = n|Z = 2) = \frac{1}{3}$  for  $n = 4$  and  $\frac{1}{6}$  for all  $n \neq 4$ ; and  $\Pr(N = n|Z = 3) = \frac{1}{3}$  for  $n = 6$  and  $\frac{1}{6}$  for all  $n \neq 6$ . Such a specification is meant to capture the possibility that the value of the auctioned object is positively correlated with potential competition. An alternative DGP where the distribution of  $N$  is restricted to be invariant in  $Z$  produces similar results. Bidders' von-Neumann-Morgenstern utility is specified as  $u(c) \equiv \left(\frac{1+c}{10}\right)^\gamma$ . The specification in this section differs slightly from that used in Section 4.3. This is mostly due to the fact that the specified supports of private values are different across the two sections. We adopt these utility specifications so that for any given  $\gamma$  the curvature of the utility function over the respective supports of  $V$  are comparable in Sections 4.3 and 5.3. The entry costs observed by potential bidders is fixed at  $K = 1$  in all auctions. As explained in Section 5.1 above, the test does not require knowledge of the true entry cost in the DGP.

Table 3 reports equilibrium entry probabilities under this DGP for  $\beta = (2, 5)$  and  $(2, 3)$  respectively. As the theoretical model suggests, for a fixed  $\beta$ , the entry probabilities are monotonically decreasing in  $N$  and increasing in  $Z$  (due to the monotonicity of  $h(Z)$ ). Also, as the marginal effect of  $Z$  diminishes from  $\beta_2 = 5$  to  $\beta_2 = 3$ , bidders are less likely to enter the bidding stage for any given  $Z$  and  $N$ . Under the current utility specification, risk-aversion appears to result in lower entry probabilities *ceteris paribus*. Most importantly, the entry probabilities are in the interior of  $(0, 1)$ . This is essential for the full-rank condition on the coefficient matrix in (12) to hold.

Table 3(a): Equilibrium Entry Probabilities:  $\beta = (2, 5)$

	$\gamma = 0.6$			$\gamma = 1$			$\gamma = 1.4$		
	$Z = 1$	$Z = 2$	$Z = 3$	$Z = 1$	$Z = 2$	$Z = 3$	$Z = 1$	$Z = 2$	$Z = 3$
$N = 2$	0.7767	0.8440	0.8756	0.9000	0.9429	0.9600	0.9547	0.9789	0.9871
$N = 3$	0.5120	0.5881	0.6299	0.6667	0.7449	0.7851	0.7714	0.8420	0.8756
$N = 4$	0.3763	0.4419	0.4797	0.5145	0.5930	0.6365	0.6213	0.7034	0.7468
$N = 5$	0.2966	0.3524	0.3854	0.4164	0.4883	0.5297	0.5152	0.5961	0.6412
$N = 6$	0.2445	0.2927	0.3216	0.3491	0.4139	0.4520	0.4386	0.5147	0.5585

Table 3(b): Equilibrium Entry Probabilities:  $\beta = (2, 3)$ 

	$\gamma = 0.6$			$\gamma = 1$			$\gamma = 1.4$		
	$Z = 1$	$Z = 2$	$Z = 3$	$Z = 1$	$Z = 2$	$Z = 3$	$Z = 1$	$Z = 2$	$Z = 3$
$N = 2$	0.7185	0.7960	0.8349	0.8571	0.9130	0.9375	0.9267	0.9626	0.9762
$N = 3$	0.4556	0.5323	0.5768	0.6051	0.6882	0.7337	0.7119	0.7914	0.8323
$N = 4$	0.3299	0.3934	0.4319	0.4573	0.5354	0.5813	0.5590	0.6435	0.6914
$N = 5$	0.2580	0.3110	0.3438	0.3660	0.4352	0.4774	0.4568	0.5366	0.5840
$N = 6$	0.2118	0.2569	0.2852	0.3045	0.3658	0.4039	0.3852	0.4585	0.5031

We report performance of the test in simulated samples that contain  $T = 3,000$  or  $T = 5,000$  auctions with variations in  $Z$  and  $N$ . For each pair  $(\beta, \gamma)$  and a sample size  $T$ , we simulate  $S = 400$  samples.<sup>16</sup> For each simulated sample, we calculate a test statistic  $\tilde{\tau}$  and record a decision under significance levels  $\alpha \in \{5\%, 10\%, 15\%\}$  respectively, based on critical values estimated from  $B = 400$  bootstrap samples. Table 4 summarizes the performance of the test for  $\gamma \in \{0.6, 0.7, 0.8, 1.0, 1.2, 1.3, 1.4\}$ .

Each row in Table 4 corresponds to a DGP with a pair  $(\beta, \gamma)$  and a sample size  $T$ . The numbers in each cell are the proportions of  $S$  simulated samples in which  $H_L, H_0, H_A$  are accepted respectively.

Table 4 (a) reports the test results when the auction heterogeneity  $Z$  has a larger marginal impact on bidders' values ( $\beta_2 = 5$ ). With a moderate sample size  $T = 3,000$ , the probabilities for rejecting the null is reasonably close to the targeted significance levels  $\alpha$  under the null. Errors in rejection probabilities also decrease as the sample size increases to  $T = 5,000$ . For both sample sizes and all significance levels, the power of the test are reasonably high for almost all alternatives except  $\gamma = 0.8$ . Besides, the power is also shown to approach 1 as sample sizes increase. 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.

---

<sup>16</sup>Note we use larger sample sizes  $T$  compared to those in Section 4.3. This is because we need sufficient observations given each pair  $(z, n)$  in the estimation of entry probabilities. (Recall that the test in Section 4.3 conditions on the number of potential bidders )

Table 4(a): Probabilities for Accepting  $[H_L, H_0, H_A]$ :  $\beta = (2, 5)$ 

	$\alpha = 5\%$	$\alpha = 10\%$	$\alpha = 15\%$
$T = 5,000$			
$\gamma = 0.6$	[0.00%, 0.75%, 99.25%]	[0.00%, 0.75%, 99.25%]	[0.00%, 0.75%, 99.25%]
$\gamma = 0.7$	[0.00%, 16.00%, 84.00%]	[0.00%, 8.00%, 92.00%]	[0.00%, 4.25%, 95.75%]
$\gamma = 0.8$	[0.00%, 80.75%, 19.25%]	[0.00%, 53.75%, 46.25%]	[0.00%, 37.00%, 63.00%]
$\gamma = 1$	[4.50%, 94.00%, 1.50%]	[6.25%, 89.25%, 4.50%]	[10.75%, 83.50%, 5.75%]
$\gamma = 1.2$	[56.25%, 43.75%, 0.00%]	[62.25%, 37.75%, 0.00%]	[64.00%, 36.00%, 0.00%]
$\gamma = 1.3$	[85.75%, 14.25%, 0.00%]	[89.25%, 10.75%, 0.00%]	[91.75%, 8.25%, 0.00%]
$\gamma = 1.4$	[97.25%, 2.75%, 0.00%]	[97.75%, 2.25%, 0.00%]	[98.25%, 1.75%, 0.00%]
$T = 3,000$			
$\gamma = 0.6$	[0.00%, 6.75%, 93.25%]	[0.00%, 4.25%, 95.75%]	[0.00%, 3.00%, 97.00%]
$\gamma = 0.7$	[0.00%, 38.00%, 62.00%]	[0.00%, 19.00%, 81.00%]	[0.00%, 13.50%, 86.50%]
$\gamma = 0.8$	[0.00%, 88.50%, 11.50%]	[0.00%, 77.75%, 22.25%]	[0.00%, 58.50%, 41.50%]
$\gamma = 1.0$	[5.00%, 92.75%, 2.25%]	[7.25%, 88.50%, 4.25%]	[11.50%, 82.25%, 6.25%]
$\gamma = 1.2$	[48.50%, 51.50%, 0.00%]	[54.50%, 45.50%, 0.00%]	[58.00%, 42.00%, 0.00%]
$\gamma = 1.3$	[76.75%, 23.25%, 0.00%]	[80.00%, 20.00%, 0.00%]	[83.00%, 17.00%, 0.00%]
$\gamma = 1.4$	[95.25%, 4.75%, 0.00%]	[97.00%, 3.00%, 0.00%]	[98.25%, 1.75%, 0.00%]

Our empirical estimates for the power of the test are low against the risk-averse alternative  $\gamma = 0.8$ . For example, when the significance level is  $\alpha = 5\%$  in Table 4(a), our empirical estimates for the power is 11.5% with  $T = 3,000$  and 19.25% with  $T = 5,000$ . However, the power improves at a fast pace as the alternative moves farther away from the null. For example, in the same Table 4(a), when  $\gamma = 0.7$  and  $\alpha = 5\%$ , the estimates for power under  $T = 3,000$  and  $5,000$  increase substantially to 62% and 84% respectively. Also, for  $\gamma$ 's with the same distance from the null value 1, the power reported against risk-loving alternatives are considerably higher than those against risk-averse alternatives. For instance, in Table 4(a), when  $\alpha = 5\%$  and  $\gamma = 1.2$ , the power is 48.5% for  $T = 3,000$  and 56.25% for  $T = 5,000$ , which are a lot greater than their counterparts under  $\gamma = 0.8$ .

That the reported power can be low against the alternatives with  $\gamma \in [0.8, 1)$  is due to the combination of two factors. First, as mentioned in Section 4.3, these risk-averse alternatives correspond to utility functions with curvatures not too far from that for a linear function. Thus the difference between the ratios compared in Proposition 5 (i.e. the probability limit of  $\tilde{\tau}$ , denoted  $\tau_*$ ) is close to zero under these DGPs. On the other hand, the standard error for estimating conditional entry probabilities are large relative to the absolute value of this difference  $\tau_*$  because of smaller sample sizes after  $Z$  and  $N$  are controlled for. (Recall the  $T$  observations consists of auctions with 15 possible combinations of  $(Z, N)$  in the DGP considered.) This in turn translates into relatively large standard errors for  $\tilde{\tau}$ . To sum up,

the test statistics are asymptotically normal both under  $\gamma = 1$  and  $\gamma = 0.8$ , with variances large relative to the difference between the means. This could explain the observed lower power against alternatives  $\gamma \in [0.8, 1)$ . Similar to the case with Section 4.3, we argue this should not be interpreted as evidence of unsatisfactory finite sample performance of our test. Rather it is due to the fact that the curvature of utility functions are close to being linear for  $\gamma \in [0.8, 1)$ . To reiterate, the power of the test does improve substantially either as the sample size further increases, or as  $\gamma$  moves farther away from 1.

Table 4(b): Probabilities for Accepting  $[H_L, H_0, H_A]$ :  $\beta = (2, 3)$

	$\alpha = 5\%$	$\alpha = 10\%$	$\alpha = 15\%$
$T = 5,000$			
$\gamma = 0.6$	[0.00%,10.00%,90.00%]	[0.00%, 6.00%, 94.00%]	[0.00%, 4.25%, 95.75%]
$\gamma = 0.7$	[0.00%,47.25%,52.75%]	[0.00%,30.25%,69.75%]	[0.00%,23.00%,77.00%]
$\gamma = 0.8$	[0.00%,88.25%,11.75%]	[0.00%,79.00%,21.00%]	[0.00%,66.75%,33.25%]
$\gamma = 1.0$	[3.00%, 94.25%, 2.75%]	[5.75%, 89.50%, 4.75%]	[9.75%, 84.00%, 6.25%]
$\gamma = 1.2$	[39.25%,60.75%,0.00%]	[45.25%,54.75%,0.00%]	[50.50%,49.50%,0.00%]
$\gamma = 1.3$	[68.25%,31.75%,0.00%]	[72.75%,27.25%,0.00%]	[77.00%,23.00%,0.00%]
$\gamma = 1.4$	[87.25%,12.75%,0.00%]	[90.50%, 9.50%, 0.00%]	[93.00%, 7.00%, 0.00%]
$T = 3,000$			
$\gamma = 0.6$	[0.00%,15.25%,84.75%]	[0.00%, 8.75%, 91.25%]	[0.00%, 6.25%, 93.75%]
$\gamma = 0.7$	[0.00%,61.25%,38.75%]	[0.00%,41.25%,58.75%]	[0.00%,31.00%,69.00%]
$\gamma = 0.8$	[0.00%, 91.25%, 8.75%]	[0.00%,84.75%,15.25%]	[0.00%,77.75%,22.25%]
$\gamma = 1.0$	[4.75%, 93.00%, 2.25%]	[8.00%, 87.50%, 4.50%]	[10.75%,83.00%,6.25%]
$\gamma = 1.2$	[33.00%,67.00%,0.00%]	[41.00%,59.00%,0.00%]	[45.25%,54.75%,0.00%]
$\gamma = 1.3$	[62.75%,37.25%,0.00%]	[67.50%,32.50%,0.00%]	[73.50%,26.50%,0.00%]
$\gamma = 1.4$	[75.75%,24.25%,0.00%]	[80.75%,19.25%,0.00%]	[84.50%,15.50%,0.00%]

Table 4(b) reports the test performance when  $\beta = (2, 3)$ . Overall, it registers the same patterns as those discussed above for Table 4(a). More interestingly, a comparison between Table 4(a) and 4(b) suggests that, for any  $T$ ,  $\alpha$  and a fixed alternative  $\gamma \neq 1$ , the power of the test is larger in the DGP with a greater marginal effect of auction heterogeneity. Such a pattern across the two panels is aligned with the idea that underlies our test: The identification of risk attitudes is driven by the difference between the rate of changes in interim utilities and the rate of changes in expected private values as the auction characteristic  $Z$  vary exogenously. The distance between these two ratios under any alternative depends on the difference in expected private values as  $Z$  varies, which depends on the marginal effect of  $Z$  on  $h(\cdot)$ .

## 5.4 Further Discussions

### 5.4.1 Endogeneity in the number of potential bidders

It is possible to extend the results from Sections 5.1-5.2 to the case where the number of potential bidders  $N$  are endogenous. The main idea is to use additional auction heterogeneities reported in data as instrumental variables under certain exclusion restrictions. Guerre, Perrigne and Vuong (2009) introduced an instrument-based argument to recover bidders' utility functions in first-price auctions. Below we discuss how their argument can be applied in the current context.

Let the number of potential bidders be given by a structural equation  $N = \iota(Z, X, \epsilon)$ , where  $(Z, X)$  are auction heterogeneities reported in data; and  $\epsilon$  is some auction-level variable not reported in data. As before, we maintain  $V_i = h(Z) + \eta_i$  for all  $i$ . Now assume the idiosyncratic component  $\eta_i$  is independent from  $(X, Z)$  conditional on  $\epsilon$ . Across potential bidders,  $\eta_i$  are independent draws from the same distribution  $F_{\eta_i|\epsilon}$  with  $E[\eta_i|\epsilon] = 0$ .

Under these conditions, the number of potential bidders may remain endogenous even after conditioning on  $(Z, X)$ . This is because these conditions allow for dependence between  $\epsilon$  and  $\eta_i$  even after conditioning on  $(Z, X)$ . On the other hand, these conditions imply that  $\eta_i$  are orthogonal to both  $(Z, N)$  conditional on  $\epsilon$ . Note  $X$  differs from  $Z$  in that  $X$  only affects potential competition but not bidders' private values. The existing literature abound in examples of such variables. In e-bay coin auctions studied by Bajari and Hortacsu (2003), sellers' reputation affects bidders' participation decisions but not their values. In U.S. Forest Service auctions considered in Athey, Levin, and Seira (2004), road costs and density of timber were found to affect participation but not bidders' values.<sup>17</sup>

Following Guerre, Perrigne and Vuong (2009), assume  $\epsilon = N - E[N|Z, X]$ . That is,  $\epsilon$  can be recovered from data as the residual of the nonparametric regression of  $N$  on  $(Z, X)$ . Our identification arguments from Section 5.1 apply after conditioning on  $\epsilon$ . Specifically, the interim utilities  $\psi_a(z, \epsilon)$  and the entry probabilities  $\rho_{a,z,n}(\epsilon)$  on the right-hand side of (9) now all depend on  $\epsilon$ , whose realized values can be recovered through regressions under our assumption above, and therefore can be controlled for. Identification of risk attitudes then follows from the same arguments as in Proposition 5 after conditioning on  $\epsilon$ .

### 5.4.2 Full identification of the utility function

We conclude this section with brief discussions about the possibility to use auction heterogeneities observed from data to fully recover bidders' utility functions. Consider a cross-sectional data with a large number of independent auctions, which share the same entry costs

---

<sup>17</sup>To be precise, the exclusion restriction in these papers are stated for entry decisions as opposed to decisions to become a *potential* bidder. Nevertheless, the arguments used in those papers could also be applies here to justify the use of these variables in  $W$  as defined above.



$k$  (unknown to econometricians) and the same number of potential bidders  $N$ , but differ in their auction-level heterogeneities  $Z$ . Then in the mixed-strategy Nash equilibrium in the entry stage,

$$\int u(y - k) dF_{Y|Z,N}(y|z) = u(0) \quad (16)$$

where  $Y_i \equiv (V_i - P_i)_+$  and  $F_{Y|Z,N}$  denotes its distribution given  $(Z, N)$ . The distribution of  $Y$  depends on  $Z$  both directly through  $F_{V_i|Z}$  and indirectly through equilibrium entry probabilities, which affect the distribution of  $P_i$  through the distribution of the number of entrants  $A$ . Note for a given  $k$ , (16) is an integral equation in  $u(\cdot)$ , with the kernel being the conditional density  $f_{Y|Z,N}$  which is identified by arguments in Proposition 1 in Section 4.1. Besides, with assumed exogenous variation in  $Z$  and  $N$ , we can augment a system of integral equations in  $u(\cdot)$  by including additional equations similar to (16) that are derived for different numbers of potential bidders and  $Z$ . We conjecture it is possible to establish uniqueness of solution in  $u(\cdot)$  and  $k$  in such a system of integral equations, after imposing location and scale normalizations (such as  $u(0) = 0$  and  $u'(0) = 1$ ) and additional primitive conditions that restricts how  $F_{Y|Z,N}$  varies with  $(Z, N)$ . We leave this as a direction for future research.<sup>18</sup>

## 6 Extensions

### 6.1 Affiliated Private Values

The assumption of independent private values is instrumental for recovering the distribution of private values from the distribution of transaction prices, for it guarantees a one-to-one mapping between a parent distribution and the distribution of the second-largest order statistic it generates. However, independence between private values fails in certain situations, e.g. if bidders' values are correlated through auction heterogeneities that bidders observe but are not reported in data. With affiliated private values, the joint distribution  $F_{(V_i)_{i \leq n}|k}$  is no longer point identified for a given  $N$  (see Athey and Haile (2002)). Nevertheless, Aradillas-Lopez, Gandhi and Quint (2012) constructs sharp bounds on bidders' ex ante expected surplus, using exogenous variation in the number of active bidders (see their Theorem 1).

---

<sup>18</sup>Let  $\Gamma(u)(z) \equiv \int u(t - k) dF_{Y|Z,N}(t|z)$ . That is,  $\Gamma$  is a linear operator mapping from the space of continuous, bounded functions over the support of  $Y$  into the space of continuous, bounded functions over the support of  $Z$ . If the model is correctly specified (i.e.  $F_{Y|Z,N}$  recovered from data is indeed generated by our model), the null space of the linear operator is necessarily a non-singleton. This is because both the truth  $u(\cdot)$  in the DGP with  $u'(\cdot) > 0$  and a constant function  $\tilde{u}$  with  $\tilde{u}(s) \equiv u(0)$  for all  $s$  are solutions of the integral equation (16). Hence one would at least need to impose further restrictions on the space of solutions in order to ensure uniqueness of the solutions in  $u(\cdot)$ .

In our model, exogenous variation in the number of entrants  $A$  follows from a more primitive condition of exogenous variation in the number of potential bidders  $N$  and from the fact that mixed-strategy entry equilibria are non-selective. As a result, we can apply the partial identification result from Aradillas-Lopez, Gandhi and Quint (2012) to construct bounds on bidders' *ex ante* expected surplus from entry. Comparing these bounds with entry costs reported in data may reveal some information about risk preference.

To see this, consider a model with no binding reserve price. Fix the number of potential bidders at  $n$  and the entry cost at  $k$ . If the joint distribution  $F_{(V_i)_{i \leq n}|k}$  are affiliated and exchangeable in bidders' indices, then there exists symmetric mixed-strategy Nash equilibria in entry stage where bidders enter independently with probability  $\lambda_k^*$  as characterized in Lemma 1 (see Lemma A2 in Appendix A for details). In addition, we maintain the following assumption about exogenous variation in the number of *potential* bidders. Let  $\bar{n}$  denote the largest integer on the support of the number of potential bidders  $N$  in the data-generating process.

**Assumption 1'** *For any  $k$ , the random vector  $(V_i)_{i \leq \bar{n}}$  is affiliated and its joint distribution is continuous and exchangeable in bidders' identities over the support  $[\underline{v}, \bar{v}]^{\bar{n}}$ . For any realized number of potential bidders  $n \leq \bar{n}$ , the joint distribution  $F_{(V_i)_{i \leq n}|k}$  equals the corresponding marginal distribution derived from  $F_{(V_i)_{i \leq \bar{n}}|k}$ .*

This assumption, together with the non-selective mixed-strategy entry equilibrium (which means the joint distribution  $F_{(V_i)_{i \in A}|k}$  equals the corresponding marginal distribution derived from  $F_{(V_i)_{i \leq n}|k}$  for any  $a \leq n$ ), suggests that the variation in the number of entrants  $A$  is exogenous (in the sense that it is not correlated with the distribution of private values). Thus the approach in Aradillas-Lopez, Gandhi and Quint (2012) can be applied to bound the *ex ante* expected profits from entry. Let  $\varphi_a(k)$  denote the expected surplus  $(V_i - P_i)_+$  for  $i$  conditioning on competing with  $A_{-i} = a$  entrants and entry cost  $k$ .

**Proposition 6** *(An Application of Theorem 1 in Aradillas-Lopez, Gandhi and Quint (2012).)*

*Under Assumption 1',  $\varphi_a^L(k) \leq \varphi_a(k) \leq \varphi_a^U(k)$  for any  $k$  and  $0 \leq a < \bar{n} - 1$ , where*

$$\begin{aligned} \varphi_a^L(k) &\equiv \frac{1}{a+1} \left( \int_{\underline{v}}^{\bar{v}} t dF_{V^{(a+1:a+1)}|k, a+1}^+(t) - \int_{\underline{v}}^{\bar{v}} t dF_{W|k, a+1}(t) \right) \text{ and} \\ \varphi_a^U(k) &\equiv \frac{1}{a+1} \left( \int_{\underline{v}}^{\bar{v}} t dF_{V^{(a+1:a+1)}|k, a+1}^-(t) - \int_{\underline{v}}^{\bar{v}} t dF_{W|k, a+1}(t) \right), \end{aligned}$$

where  $F_{W|k, s}$  is the distribution of transaction price given  $k$  and  $A = s$ , and

$$\begin{aligned} F_{V^{(s:s)}|k, s}^+(t) &= \sum_{m=s+1}^{\bar{n}} \frac{s}{m(m-1)} F_{V^{(m-1:m)}|k}(t) + \frac{s}{\bar{n}} F_{V^{(\bar{n}-1:\bar{n})}|k}(t) \\ F_{V^{(s:s)}|k, s}^-(t) &= \sum_{m=s+1}^{\bar{n}} \frac{s}{m(m-1)} F_{V^{(m-1:m)}|k}(t) + \frac{s}{\bar{n}} \left[ \phi_{\bar{n}}(F_{V^{(\bar{n}-1:\bar{n})}|k}(t)) \right]^{\bar{n}}. \end{aligned}$$

The intuition of this result is as follows. First, an application using the Law of Total Probability reveals:

$$F_{V^{(a+1:a+1)}|k} = \frac{1}{a+2} F_{V^{(a+1:a+2)}|k} + \frac{a+1}{a+2} F_{V^{(a+2:a+2)}|k}.$$

Then recursive substitutions show that  $F_{V^{(a+1:a+1)}|k}$  (which is not point-identified from transaction price distributions due to affiliation) can be written as a linear combination of distributions of second-largest order statistic  $F_{V^{(m-1:m)}|k}$  for  $m = a+2, \dots, \bar{n}$  (which is directly identifiable from the data) and the distribution of the largest order statistic  $F_{V^{(\bar{n}:\bar{n})}|k}$ . Second, it can be shown that the distribution of  $V^{(\bar{n}:\bar{n})}$  given  $k$  is bounded between  $[\phi_{\bar{n}}^{-1}(F_{V^{(\bar{n}-1:\bar{n})}|k}(\cdot))]^{\bar{n}}$  and  $F_{V^{(\bar{n}-1:\bar{n})}|k}(\cdot)$ . These bounds are derived by exploiting the link between the distribution of first- and the second-order statistics of  $V_i$  (i.e.  $F_{V^{(\bar{n},\bar{n})}|k}$  and  $F_{V^{(\bar{n}-1:\bar{n})}|k}$ ) under two extreme scenarios: full independence or perfect correlation among private values. These two results imply that  $F_{V^{(a:a)}|k}$  is bounded between  $F_{V^{(a:a)}|k,a}^+$  and  $F_{V^{(a:a)}|k,a}^-$  for all  $a \leq \bar{n} - 1$ . Proposition 6 then follows from the fact that the ex ante surplus  $\varphi_a(k)$  can be expressed as the difference between  $\frac{1}{a+1} \int_{\underline{v}}^{\bar{v}} t dF_{V^{(a+1:a+1)}|k}(t)$  and  $\frac{1}{a+1} \int_{\underline{v}}^{\bar{v}} t dF_{V^{(a:a)}|k}(t)$ .

It then follows from Proposition 6 that a bidder's *ex ante* surplus prior to entry in the presence of  $n$  potential bidders (denoted by  $\varphi^*(k, n)$ ) is bounded between  $\varphi_L^*(k, n)$  and  $\varphi_U^*(k, n)$  for all  $n < \bar{n}$ , where

$$\varphi_h^*(k, n) \equiv \sum_{a=0}^{n-1} \varphi_a^h(k) \Pr(A_{-i} = a | k, n) \text{ for } h \in \{L, U\}. \quad (17)$$

Unlike in the case with independent private values where we could point identify  $\varphi^*(k, n)$ , here we can only recover a pair of bounds  $\varphi_L^*(k, n)$  and  $\varphi_U^*(k, n)$  when private values are affiliated. However, provided entry costs are observed from data, one can still test the chain of inequalities  $\varphi_L^*(k, n) \leq k \leq \varphi_U^*(k, n)$  against the alternatives of “ $k < \varphi_L^*(k, n)$ ” and “ $k > \varphi_U^*(k, n)$ ”, using a test statistic based on sample analogs (whose finite sample distribution can be estimated using the bootstrap). If the null is rejected in favor of one of the alternatives, we can conclude there is evidence in the data that supports hypotheses of risk-aversion or risk-loving. Unfortunately, on the other hand, a failure to reject the null does not necessarily allow us to conclude whether there is significant evidence for risk-neutrality or not.

The argument in the preceding paragraph can be extended where entry costs are only measured with noises, and are orthogonal from the joint distribution of private values. In this case, bounds on ex ante surplus  $\varphi^*(k, n)$  in (17) still hold for all  $k$  given any  $n$ , except that the bounds on interim surplus  $\{\varphi_a^h\}_{h=L,U}$  no longer depend on the unobserved cost  $k$ . Thus, similar to the case with IPV in Corollary 1, we can identify  $E[\varphi_h^*(K, n)|n]$  as  $\sum_{a=0}^{n-1} \varphi_a^h \Pr(A_{-i} = a | n)$  for  $h = L, U$ . Thus, provided the measurement errors are zero-mean so that  $E[K]$  can be consistently estimated, we can construct a consistent test for the null “ $E[\varphi_L^*(K, n)|n] \leq E[K] \leq E[\varphi_U^*(K, n)|n]$ ” against the alternatives “ $E[K] < E[\varphi_L^*(K, n)|n]$ ” and “ $E[K] > E[\varphi_U^*(K, n)|n]$ ”. As before, a rejection of the null would provide statistically significant evidence against risk-neutrality while a failure to rejection would leave the test inconclusive.

## 6.2 Selective Entry with Informative Signals

Consider a model where each potential bidder observes in the entry stage a private signal correlated with his private values. Thus entry decisions depend on these informative signals. We identify bidders' risk attitudes in such a model, assuming data report continuous variations in entry costs across auctions. As before, we fix  $N$  (the number of potential bidders) and suppress it from the notations. Auction models with endogenous and selective entry have been studied in Ye (2007), Gentry and Li (2013) and Marmer, Shneyerov and Xu (2013). Li, Lu and Zhao (2012) considers the same model of auctions with risk-averse bidders, affiliated signals and selective entry. Their emphasis is on the inference of the form of risk-aversions using data from both first-price and ascending auctions.

The model is specified as follows. In an entry stage, each bidder  $i$  observes a private signal  $S_i$  and decides whether to pay costs  $k$  and enter the bidding stage. The entry cost is commonly known and the same for all bidders. Upon entry in the bidding stage, each entrant draws a private value  $V_i$ , and competes in an ascending auction with a reserve price  $r$ . The joint distribution  $F(S_1, \dots, S_N, V_1, \dots, V_N)$  and  $r$  are common knowledge among potential bidders. Each entrant may not be aware of the number of active competitors  $A$  while bidding.

**Assumption S** (i)  $(S_i, V_i)$  are identically and independently distributed across potential bidders ( $F(S_1, \dots, S_N, V_1, \dots, V_N) = \prod_{i=1}^N F(S_i, V_i)$ ); (ii) For each  $i$ ,  $(S_i, V_i)$  are affiliated; (iii) Marginal distributions of  $S_i$  and  $V_i$ , denoted  $F_S$  and  $F_V$  respectively, are continuous and increasing over bounded support  $[\underline{s}, \bar{s}]$  and  $[\underline{v}, \bar{v}]$  respectively.

Suppose the reserve price is binding with  $r > \underline{v}$ . Let  $\bar{w}_i(s_i, k; s_{-i})$  denote ex ante utility for bidder  $i$  with signal  $s_i$  if potential competitors follow monotone, pure-strategy Bayesian Nash equilibria characterized by cutoffs  $s_{-i} \equiv (s_j)_{j \neq i}$ . That is,  $\bar{w}_i(s_i, k; s_{-i}) \equiv E[u((V_i - P_i)_+ - k) | S_i = s_i, A_{-i} = \{j \neq i : S_j \geq s_j\}, i \in A]$ .

**Lemma 3** Under Assumption S,  $\bar{w}_i(s_i, k; s_{-i})$  is increasing in  $s_i$  and non-decreasing in  $s_{-i}$  for any  $k$ .

Let  $\bar{w}(s, k)$  be a shorthand for  $\bar{w}_i(s, k; (s, \dots, s)) = E[u((V_i - P_i)_+ - k) | S_i = s, A_{-i} = \{j \neq i : S_j \geq s\}, i \in A]$ . Under Assumption S,  $\bar{w}(s, k)$  is the same for all  $i$ , and is increasing in  $s$  due to Lemma 3. An arguments similar to that in Lemma 1 show that under Assumption S, there exists the following unique pure-strategy BNE in the entry stage: A potential bidder  $i$  enters iff  $s_i \geq s_k^*$ , where  $s_k^*$  solves  $\bar{w}(s_k^*, k) = u(0)$  if  $\bar{w}(\underline{s}, k) \leq u(0) \leq \bar{w}(\bar{s}, k)$ . Otherwise,  $s_k^* = \underline{s}$  (or  $s_k^* = \bar{s}$ ) if  $\bar{w}(\underline{s}, k) > u(0)$  (or  $\bar{w}(\bar{s}, k) < u(0)$  respectively). Entry decisions in this model are selective in that the distribution of private values conditional on entry differs from their unconditional distribution.

To find bidders' risk attitudes when there is selective entry, we again exploit the fact that concave utilities lead to positive risk premia in the entry decisions. Let  $\bar{\pi}(s, k) \equiv E[(V_i - P_i)_+ - k | S_i = s, A_{-i} = \{j \neq i : S_j \geq s\}]$ , which is identical for all  $i$  under Assumption

S. By the same argument as in Section 4.1, we can construct a test for bidders' risk attitudes as long as  $\bar{\pi}(s_k^*, k)$  is identified from distributions of transaction prices and entry decisions for any  $k$  that yields non-degenerate entry probabilities in equilibrium. Specifically, consider some costs  $k$  with  $\underline{s} < s_k^* < \bar{s}$  so that uninteresting cases involving degenerate entry decisions are ruled out. (Whether a level of entry cost  $k$  leads to any non-degenerate entry probabilities is testable using the distribution of entry decisions from data.) Then, under Assumption S,  $\bar{\pi}(s_k^*, k) = 0$  when bidders are risk-neutral; and  $\bar{\pi}(s_k^*, k) > 0$ , or  $< 0$ , when bidders are risk-averse, or respectively risk-loving.

The test for risk attitudes introduced in Section 4.2 is not directly applicable under selective entry, because it leads to distorted decisions in favor of risk-aversion even when the true DGP has risk-neutral bidders. Such distortion arises from the affiliation between private values  $V_i$  and entry signals  $S_i$ . To see why, recall from the preceding paragraph that, with selective entry, bidders' risk attitudes are linked only to the sign of risk-premia for a marginal bidder whose signal equals the cutoff  $s_k^*$  that characterizes the equilibrium entry strategy. If we were to apply the test from Section 4.2 under the current context of selective entry, it would amount to testing the sign of ex ante payoffs for a generic entrant whose signal is not restricted to be equal to the cutoff (that is,  $\tilde{\pi}(s_k^*, k) \equiv E[(V_i - P_i)_+ - k | S_i \geq s_k^*, A_{-i} \equiv \{j \neq i : S_j \geq s_k^*\}]$ ). However, the affiliation between  $S_i$  and  $V_i$  suggests ex ante payoffs for other *non-marginal* entrants must be greater than the marginal entrant's risk premium  $\bar{\pi}(s_k^*, k)$ . That is,  $\tilde{\pi}(s_k^*, k)$  is bounded below by  $\bar{\pi}(s_k^*, k)$ . Consequently, applying the test from Section 4.2 for models with selective entry leads to over-rejection of risk-neutrality in favor of risk-aversion.<sup>19</sup>

The next proposition shows how to recover  $\bar{\pi}(s_k^*, k)$  from entry probabilities and the distribution of transaction prices. To do this, it suffices to recover the distribution of  $V_i$  given  $S_i = s_k^*$  and the distribution of  $P_i$  given  $A_{-i} = \{j \neq i : S_j \geq s_k^*\}$  from data.

**Proposition 7** *Suppose (i) Assumption S holds with  $(V_i, S_i)_{i \in N}$  independent from entry costs; (ii) for some fixed  $k$ , there exists  $\varepsilon > 0$  such that  $\underline{s} < s_{k'}^* < \bar{s}$  for all  $k' \in (k - \varepsilon, k + \varepsilon)$ ; and (iii) entry costs are observed in the data. Then  $\bar{\pi}(s_k^*, k)$  is identified from the distribution of transaction prices and entry probabilities.*

Implementing a test based on Proposition 7 would require the researcher to observe a continuous variation of entry costs known to potential bidders. As discussed in Section 4, such an assumption is strong and has limitations. It could fail under various empirical environments (say due to bias in the measurement of entry costs). It remains an open question whether additional exogenous variations could be exploited to construct a test under affiliated endogenous entry without requiring such knowledge. We also leave the definition of a test statistic based on Proposition 7 and its asymptotic properties to future research.

---

<sup>19</sup>We are grateful to a referee for this insight.

## 7 Conclusion

We propose tests for bidders' risk attitudes in ascending auctions where potential bidders make endogenous entry decisions based on their information of entry costs and *ex ante* knowledge of the distribution of private values. First, we show that risk premium can be non-parametrically recovered from the distribution of transaction prices and entry decisions, as long as the expected entry cost is identified from the data. A test is proposed and shown to have good finite sample performance under various data-generating processes. Second, we show how exogenous variations in number of potential bidders and observed changes in auction features could provide additional identifying power to infer risk attitudes even when knowledge about entry costs is not available. Finally, we extend our results to identify bidders' risk preference in a more general model where entries into auctions are selective. Discussions about possible extensions to cases with affiliated private values are also provided.

There are several directions for future research. First, is it possible to recover the full utility functions, not just the sign of risk attitudes? We conjecture this is possible with additional variations in entry costs or auction heterogeneities reported in data. Some additional restrictions on how these factors affect private values may be needed to complete this task. Second, can we extend the analysis to ascending auctions with discrete increments? Haile and Tamer (2003) show how to form sharp bounds on value distributions using price distributions in this case. These bounds could in turn lead to bounds on expected risk premia. Thus a test could be constructed to test the null that zero is covered by these bounds on risk premia. Similar to the case with affiliated private values in Section 6.1, rejection of the null could lead to a decision on risk attitudes while failure in rejections leaves the test inconclusive. It remains an open question how to construct more informative tests under additional economic restrictions on the model structure.

# APPENDIX

## A Proofs of Identification

**Proof of Lemma 1.** We use the following property of expected utility from entry.

**Lemma A1** Under Assumption 1,  $\omega(k; \lambda_{-i})$  is continuous and decreasing in  $\lambda_{-i}$  for all  $k$ .

*Proof of Lemma A1.* By Law of Iterated Expectations,

$$\omega(k; \lambda_{-i}) = u(-k)F_{V|k}(r) + \int_r^{\bar{v}} h(v, k; \lambda_{-i})dF_{V|k}(v) \quad (\text{A1})$$

where for  $v > r$ ,

$$\begin{aligned} h(v, k; \lambda_{-i}) &\equiv u(v - r - k)F_{P_i}(r|k, \lambda_{-i}) + \int_r^v u(v - p - k)dF_{P_i}(p|k, \lambda_{-i}) \\ &+ u(-k)[1 - F_{P_i}(v|k, \lambda_{-i})] \end{aligned}$$

and we have used the independence between private values  $V_i$  conditional on entry costs. Note for any  $t \in [r, \bar{v}]$ , the event " $P_i \leq t$ " can be represented as

$$\bigcap_{j \in N \setminus \{i\}} \{ "j \text{ stays out}" \text{ or } "j \text{ enters} \cap V_j \leq t" \}.$$

Due to the independence between entry decisions and between private values,  $F_{P_i}(t|k, \lambda_{-i}) = \prod_{j \neq i} [1 - \lambda_j + \lambda_j F_{V|k}(t)]$ . The marginal effect of  $\lambda_j$  on this conditional probability is strictly negative for all  $j \neq i$  at  $\lambda_j \in [0, 1]$  and  $t \in [r, \bar{v}]$ . (Recall  $F_{V|k}(r) > 0$  when  $r$  is binding.) This implies  $h(v, k; \lambda_{-i})$  is decreasing in  $\lambda_{-i}$  for any  $k$ . Hence  $\omega(k; \lambda_{-i})$  is decreasing in  $\lambda_{-i}$ . Continuity follows from Dominated Convergence Theorem. *Q.E.D.*

The rest of the proof is similar to the risk-neutral case in Levin and Smith (1994) and omitted.  $\square$

**Proof of Proposition 1.** Conditional on  $k$ , entry decisions are independent across bidders, and jointly independent from private values. Besides, private values are i.i.d. across bidders given  $k$ . Hence, once conditional on  $k$  and the realization of  $A_{-i}$ ,  $(V_i, P_i)$  are independent from mixed strategies adopted by potential competitors. Thus

$$E[(V_i - P_i)_+ | k] = \sum_{a=0}^{N-1} E[(V_i - P_i)_+ | k, A_{-i} = a] \Pr(A_{-i} = a | k). \quad (\text{A2})$$

With  $k$  and entry decisions observed from data,  $\lambda_k^*$  is directly identified as the probability that a bidder enters under cost  $k$ . Consequently,  $\Pr(A_{-i} = a | k)$  can be recovered as a binomial distribution with parameters  $N - 1$  and  $\lambda_k^*$ . Conditional on entering with cost  $k$ ,

private values are independent draws from  $F_{V|k}$ . Let  $1\{\cdot\}$  denote the indicator function. By Law of Iterated Expectations,  $E[(V_i - P_i)_+ | k, A_{-i} = a]$  is

$$\begin{aligned} & E[(V_i - P_i)1\{V_i > P_i > r\} | k, A_{-i} = a] + E[(V_i - r)1\{V_i > r\}1\{P_i = r\} | k, A_{-i} = a] \\ &= \int_r^{\bar{v}} \left( \int_v^{\bar{v}} (s - v) dF_{V|k}(s) \right) dF_{V|k}(v)^a + F_{V|k}(r)^a \int_r^{\bar{v}} (v - r) dF_{V|k}(v) \end{aligned} \quad (\text{A3})$$

for all  $A_{-i} = a \geq 1$ , due to independence between  $V_i$  and  $P_i$  given  $k$ . Applying integration by parts to the first term on the right-hand side of (A3), we have

$$E[(V_i - P_i)_+ | A_{-i} = a, k] = \int_r^{\bar{v}} (F_{V|k}(v)^a - F_{V|k}(v)^{a+1}) dv$$

for  $a \geq 1$ . Besides,  $E[(V_i - P_i)_+ | A_{-i} = 0, k] = E[(V_i - r)_+ | k] = \int_r^{\bar{v}} [1 - F_{V|k}(v)] dv$ .

With  $k$  observed from data,  $\pi(k)$  is recoverable as long as  $F_{V|k}(v)$  is identified for  $v \geq r$ . Let  $W$  denote transaction prices observed in data. If no entrants bid above  $r$ , then define  $W < r$ . The symmetric IPV assumption implies for any  $m \geq 2$ ,  $\Pr(W < r | A = m, k) = \Pr(V^{(m:m)} < r | k) = F_{V|k}(r)^m$  and  $\Pr(W = r | A = m, k) = mF_{V|k}(r)^{m-1}[1 - F_{V|k}(r)]$ . Hence for any  $m \geq 2$  and  $t \geq r$ ,

$$\begin{aligned} \Pr(W \leq t | A = m, k) &= \Pr(W < r | m, k) + \Pr(W = r | m, k) + \Pr(r < W \leq t | m, k) \\ &= F_{V|k}(t)^m + mF_{V|k}(t)^{m-1}[1 - F_{V|k}(t)] = F_{V^{(m-1:m)}|k}(t). \end{aligned}$$

For any  $m \geq 2$ , define  $\phi_m(t) \equiv t^m + mt^{m-1}(1-t)$  so that  $F_{V^{(m-1:m)}|k}(t) = \phi_m(F_{V|k}(t))$ . Since  $\phi_m(t)$  is one-to-one for any  $m \geq 2$  over  $t \in [0, 1]$ ,  $F_{V|k}(t)$  is (over-)identified for each  $t \geq r$  from the distributions of  $W$  conditional on  $k$  and  $A = m$ .  $\square$

**Proof of Corollary 1.** *Part (a).* By Proposition 1 and the condition in part (a),  $\pi(k) = 0$  for all  $k \in [\underline{k}, \bar{k}]$  if bidders are risk-neutral, and  $\pi(k) > 0$  (or  $\pi(k) < 0$ ) for all  $k$  if bidders are risk-averse (or risk-loving). Integrating out  $k$  using  $F_K$  proves (a).

*Part (b).* Independence between  $K$  and  $(V_i)_{i \in N}$  implies that, given  $A_{-i} = a$ , the vector of order sorted values  $\{V^{(s:a+1)}\}_{s=1}^{a+1}$  is independent from  $K$ . Thus  $\varphi_a \equiv E[(V_i - P_i)_+ | A_{-i} = a, k]$  does not depend on  $k$  for all  $a \geq 0$  (recall  $P_i \equiv r$  when  $A_{-i} = \emptyset$ ). By Law of Total Probability,  $E[\pi(K)]$  is:

$$\int_{\underline{k}}^{\bar{k}} \left( -k + \sum_{a=0}^{N-1} \varphi_a \Pr(A_{-i} = a | k) \right) dF_K(k) = \sum_{a=0}^{N-1} \varphi_a \Pr(A_{-i} = a) - \mu_K. \quad (\text{A4})$$

To identify  $\Pr(A_{-i} = a)$  (or  $\int_{\underline{k}}^{\bar{k}} \Pr(A_{-i} = a | k) dF_K(k)$ ), note that given any  $k$  and  $N$ ,  $A_{-i}$  is binomial  $(N-1, \lambda_k^*)$  while  $A$  is binomial  $(N, \lambda_k^*)$ . By construction,

$$\Pr(A_{-i} = a | k) = \frac{N-a}{N} \Pr(A = a | k) + \frac{a+1}{N} \Pr(A = a+1 | k) \quad (\text{A5})$$



for all  $k$  and  $0 \leq a \leq N - 1$ . Integrating out  $k$  on both sides of (A5) implies  $\Pr(A_{-i} = a) = \frac{N-a}{N} \Pr(A = a) + \frac{a+1}{N} \Pr(A = a + 1)$ . Since the unconditional distribution of  $A$  is directly identified, so is the distribution of  $A_{-i}$ .<sup>20</sup> With  $\mu_K$  assumed known, this implies  $E[\pi(K)]$  is identified.  $\square$

**Proof of Lemma 2.** First note  $\psi_1^* \neq u(0)$  under Assumption 3. [To see this, consider the case with  $z$  and  $N = 2$ . Then  $\rho_{0,z,2}\psi_0(z) + \rho_{1,z,2}\psi_1^* = u(0)$  with  $\rho_{0,z,2} > 0$ ,  $\rho_{1,z,2} > 0$  and  $\rho_{0,z,2} + \rho_{1,z,2} = 1$ . It follows from Assumption 2 that  $\psi_0(z) > \psi_1^*$ . This implies  $\psi_1^* < u(0)$ .] By construction, the  $r$ -th row of the coefficient matrix on the left-hand side of (12) and the  $r$ -th coordinate of the coefficient vector in front of  $\psi_1^*$  on the right-hand side of (12) add up to 1. Hence we can write (12) as

$$\overline{\mathbf{M}} \cdot [(\psi_0(z''), \psi_0(z'), \psi_0(z), \psi_2^*, \dots, \psi_{n-1}^*)' - u(0)] = -\overline{\boldsymbol{\rho}}_1 \cdot [\psi_1^* - u(0)]. \quad (\text{A6})$$

where  $\overline{\mathbf{M}}$  denotes the  $(n + 1)$ -by- $(n + 1)$  matrix of coefficients on the left-hand side of (12); and  $\overline{\boldsymbol{\rho}}_1$  denotes the  $(n + 1)$ -by-1 vector of coefficients in front of  $\psi_1^*$  on the right-hand side of (12).

Next, we show that the matrix of coefficients  $\overline{\mathbf{M}}$  has full-rank  $n + 1$ . To do this, it suffices to show each one of the  $n + 1$  rows cannot be a linear combination of the other  $n$  rows. Under Assumption 3,  $\lambda_{0,z'',n} > 0$  and  $\lambda_{0,z',n} > 0$ , so neither of the first two rows is a linear combination of the other  $n$  rows. As for the third row, it cannot be a linear combination of the other  $n$  rows, with nonzero weights for the first two rows that correspond to  $z'$  and  $z''$ . Next, suppose it were a linear combination with nonzero weights for the last  $n - 2$  rows of the coefficient matrix in (12) only. Then this would contradict the assumption that the  $(n - 1)$ -by- $(n - 1)$  block in the lower right corner of the coefficient matrix is a lower triangular matrix with all entries below (and including) the diagonal vector being positive.

With  $\psi_1^* < u(0)$  and the coefficient matrix  $\overline{\mathbf{M}}$  being non-singular, (A6) can be rescaled and written as

$$\overline{\mathbf{M}} \cdot (\tilde{\psi}_{0,z''}, \tilde{\psi}_{0,z'}, \tilde{\psi}_{0,z}, \tilde{\psi}_2, \dots, \tilde{\psi}_{n-1})' = -\overline{\boldsymbol{\rho}}_1 \cdot c \quad (\text{A7})$$

where

$$(\tilde{\psi}_{0,z''}, \tilde{\psi}_{0,z'}, \tilde{\psi}_{0,z}, \tilde{\psi}_2, \dots, \tilde{\psi}_{n-1})' \equiv \frac{c}{\psi_1^* - u(0)} [(\psi_0(z''), \psi_0(z'), \psi_0(z), \psi_2^*, \dots, \psi_{n-1}^*)' - u(0)].$$

The lemma then follows immediately.  $\square$

**Proof of Proposition 5.** *Part (a).* It suffices to note that, under Assumption 2,  $E(W|Z = z, A = a) = h(z) + E[\eta^{(a-1:a)}]$  for all  $z$  and  $a \geq 2$ , and  $E[\eta^{(a-1:a)}]$  is independent from  $z$ .

---

<sup>20</sup>In fact this provides us with a set of testable implications that can be used for testing the model specification of symmetric IPV.

Part (b). By construction,

$$\psi_0(z) \equiv E[u(V_i - k)|z] = E[u(h(z) + \eta_i - k)]$$

where the first expectation is taken with respect to  $V_i$  given  $Z = z$  and the second is with respect to  $\eta_i$  alone while  $z$  is fixed at a realized value. By Lemma 2,  $(\tilde{\psi}_{0,z'} - \tilde{\psi}_{0,z})/(\tilde{\psi}_{0,z''} - \tilde{\psi}_{0,z'})$  equals  $[\psi_0(z') - \psi_0(z)]/[\psi_0(z'') - \psi_0(z)]$ , which in turn equals:

$$\frac{E[u(h(z') + \eta_i - k)] - E[u(h(z) + \eta_i - k)]}{E[u(h(z'') + \eta_i - k)] - E[u(h(z') + \eta_i - k)]}. \quad (\text{A8})$$

By independence between  $Z$  and  $\eta$  in Assumption 2 and the Mean Value Theorem, (A8) equals:

$$\frac{[h(z') - h(z)] E[u'(\alpha h(z') + (1 - \alpha)h(z) + \eta_i - k)]}{[h(z'') - h(z')] E[u'(\alpha' h(z'') + (1 - \alpha')h(z') + \eta_i - k)]} \quad (\text{A9})$$

for some  $\alpha$  and  $\alpha' \in (0, 1)$  that depend on  $(z, z')$  and  $(z', z'')$  respectively. By the ordering of  $h(z'')$ ,  $h(z')$  and  $h(z)$ , it must be the case that

$$\alpha h(z') + (1 - \alpha)h(z) + \eta_i - k < \alpha' h(z'') + (1 - \alpha')h(z') + \eta_i - k$$

for any  $k$ ,  $\eta_i$ ,  $\alpha$  and  $\alpha'$ .

If bidders are risk-neutral, then

$$u'(\alpha h(z') + (1 - \alpha)h(z) + \eta_i - k) = u'(\alpha' h(z'') + (1 - \alpha')h(z') + \eta_i - k) > 0 \quad (\text{A10})$$

for all  $k$ ,  $\eta_i$ ,  $\alpha$  and  $\alpha'$ , and (A9) equals  $[h(z') - h(z)]/[h(z'') - h(z')]$ . Otherwise if bidders are risk-averse then the equality in (A10) is replaced by a strict inequality “ $>$ ” for all  $k$ ,  $\eta_i$ ,  $\alpha$  and  $\alpha'$  due to the concavity of  $u$ . It then follows from the independence between  $\eta_i$  and  $Z$  that

$$\frac{E[u'(\alpha h(z') + (1 - \alpha)h(z) + \eta_i - k)]}{E[u'(\alpha' h(z'') + (1 - \alpha')h(z') + \eta_i - k)]} > 1$$

when bidders are risk-averse. Therefore (A9) is strictly greater than  $[h(z') - h(z)]/[h(z'') - h(z')]$  under risk-aversion. Symmetric arguments show the reverse strict inequality holds when bidders are risk-loving.  $\square$

**Lemma A2** *Let the joint distribution of private values  $F_{(V_i)_{i=1}^n|k}$  be affiliated and exchangeable for a given number of potential bidders  $n$  and entry cost  $k$ . If  $\omega(k; (1, \dots, 1)) < u(0) < \omega(k; (0, \dots, 0))$  where  $\omega$  is defined as in (1), then there exists a unique symmetric BNE in which all bidders enter independently with probability  $\lambda_k^*$ , where  $\lambda_k^*$  solves  $\omega(k; (\lambda_k^*, \dots, \lambda_k^*)) = u(0)$ .*

*Proof of Lemma A2.* It suffices to show  $F_{P_i}(t|k, \boldsymbol{\lambda}_{-i})$  is decreasing in  $\boldsymbol{\lambda}_{-i}$  given  $k$  and  $n$ . To see this, note by the Law of Total Probability,

$$F_{P_i}(t|k, \boldsymbol{\lambda}_{-i}) \equiv \sum_{a=0}^{n-1} F_{V^{(a:a)}|k}(t) \Pr(A_{-i} = a|\boldsymbol{\lambda}_{-i}, k),$$

where  $F_{V^{(a:a)}|k}(t)$  denotes the distribution of the largest order statistic from a  $a$ -dimensional random vector  $(V_i)_{i \leq a}$ ; and define  $F_{V^{(0:a)}|k}(t) \equiv 1$  for  $t \geq 0$ . Under the assumptions of the lemma,  $F_{V^{(a':a')}|k}$  first-order stochastically dominates  $F_{V^{(a:a)}|k}$  if  $a' > a$ . Besides,  $\Pr(A_{-i} = a | \boldsymbol{\lambda}_{-i}, k)$  is stochastically increasing in  $\boldsymbol{\lambda}_{-i}$  since bidders enter independently with respective probabilities. Thus  $F_{P_i}(t|k, \boldsymbol{\lambda}_{-i})$  is decreasing in  $\boldsymbol{\lambda}_{-i}$  given  $k$  and  $n$ . The rest of the proof follows from the same arguments of Lemma A1.  $\square$

**Proof of Lemma 3.** Let  $P_i \equiv \max_{j \in A \setminus \{i\}} \{\max\{V_j, r\}\}$ . Under Assumption S,  $F(p_i|v_i, s_i, s_{-i}) = F(p_i|s_{-i})$  and  $F(v_i|s_i, s_{-i}) = F(v_i|s_i)$ . Thus by the Law of Iterated Expectations,

$$\bar{\omega}_i(s_i, k; s_{-i}) = u(-k) \Pr(V_i \leq r | s_i) + \int_r^{\bar{v}} \bar{h}(v, k; s_{-i}) dF_{V_i|S_i}(v|s_i)$$

where  $\bar{h}(v, k; s_{-i}) \equiv u(v - r - k)F_{P_i|S_{-i}}(r|s_{-i}) + \int_r^v u(v - p - k) dF_{P_i|S_{-i}}(p|s_{-i}) + u(-k)[1 - F_{P_i|S_{-i}}(v|s_{-i})]$ . By the Leibniz Rule,

$$\frac{\partial}{\partial v} \bar{h}(v, k; s_{-i}) = u'(v - r - k)F_{P_i|S_{-i}}(r|s_{-i}) + \int_r^v u'(v - p - k) dF_{P_i|S_{-i}}(p|s_{-i}) > 0 \quad (\text{A11})$$

Thus  $\bar{h}$  is increasing in  $v$  for fixed  $s_{-i}$  and  $k$ . By the affiliation of  $V_i$  and  $S_i$  for all  $i$ , the distribution  $F(\cdot|s_i)$  is stochastically increasing in  $s_i$ . Hence  $\bar{\omega}_i(s_i, k; s_{-i})$  is increasing in  $s_i$  given  $k$  and  $s_{-i}$ . To show  $\bar{\omega}_i(s_i, k; s_{-i})$  is non-decreasing in  $s_{-i}$  given  $s_i$  and  $k$ , it suffices to show  $F_{P_i|S_{-i}}(p|s_{-i})$  is stochastically non-decreasing in  $s_{-i}$  for all  $p \geq r$ , which would imply  $\bar{h}(v, k; s_{-i})$  is non-decreasing in  $s_{-i}$  for any  $v \in [\underline{v}, \bar{v}]$ . Note for any  $t \in [r, \bar{v}]$ , the event " $P_i \leq t$ " can be written as

$$\cap_{j \neq i} \{ "S_j < s_j" \cup "S_j \geq s_j \cap V_j \leq t" \}$$

Under Assumption S,  $\Pr(P_i \leq t | s_{-i}) = \prod_{j \neq i} [F_{S_j}(s_j) + \Pr(V_j \leq t, S_j \geq s_j)]$ . Also note for all  $t$  and any  $s'_j > s_j$ ,

$$F_{S_j}(s_j) + \Pr(V_j \leq t, S_j \geq s_j) \leq F_{S_j}(s'_j) + \Pr(V_j \leq t, S_j \geq s'_j).$$

Hence  $F_{P_i|S_{-i}}(t|s_{-i})$  is non-decreasing in  $s_{-i}$  for all  $t \in [r, \bar{v}]$ .  $\square$

**Lemma A3** *Under Assumption S, for any  $k$  such that  $\underline{s} < s_k^* < \bar{s}$ ,  $F_V(t|S_i \geq s_k^*)$  is identified for  $t \geq r$  from the distribution of  $W$  conditional on any  $k$  and any number of entrants  $a$  (with  $a \geq 2$ ).*

*Proof of Lemma A3.* By definition, for any  $t \geq r$ ,  $\Pr(W \leq t | k, A = a)$  is identical to the distribution of the second-highest order statistic among  $a$  independent draws from the same conditional distribution  $F_V(\cdot | S_i \geq s_k^*)$ . That is, for any  $t \geq r$

$$\Pr(W \leq t | k, a) = \sum_{m=a-1}^a \binom{n}{m} F_V(t | S_i \geq s_k^*)^m [1 - F_V(t | S_i \geq s_k^*)]^{a-m}$$

Thus for all  $a \geq 2$ , there exists a one-to-one mapping  $\phi_a$  so that  $F_V(t|S_j \geq s_k^*) = \phi_a(\Pr(W \leq t|k, a))$ . (Note the mapping  $\phi_a$  does not depend on  $k$  as  $(V_i, S_i)$  is assumed to be independent from entry costs.) That  $\Pr(i \text{ enters}|k) > 0$  for  $k$  implies  $\Pr(A \geq 2|k) > 0$ . Thus  $F_V(t|S_i \geq s_k^*)$  is over-identified for  $t \geq r$ , because the identification arguments above can be applied for any  $a$  such that  $\Pr(A = a|k) > 0$ .  $\square$

**Lemma A4** *Under Assumption S,  $F_V(t|S_i = s_k^*)$  is identified for any  $t \geq r$  and any  $k$  s.t.  $\underline{s} < s_k^* < \bar{s}$ , provided  $\Pr(W \leq t|k, a)$  and  $\Pr(i \text{ enters}|k)$  are observed in an open neighborhood around  $k$  for some  $a \geq 2$ .*

*Proof of Lemma A4.* Given Lemma A3,  $F_V(t|S_i \geq s_k^*)$  is identified for any such  $k$  using distributions of entry decisions and transaction prices. Hence

$$\begin{aligned} \Pr(V_i \leq t, S_i \geq s_k^*) &= F_{V|S}(t|S_i \geq s_k^*) \Pr(i \text{ enters}|k) \\ &= \phi_a^{-1}(\Pr(W \leq t|k, a)) \Pr(i \text{ enters}|k) \end{aligned} \quad (\text{A12})$$

is also identified using any  $a$  such that  $\Pr(A = a|k) > 0$ . We consider this joint distribution as known for the rest of the proof. For any  $t \geq r$ , differentiating this distribution with respect to entry costs at  $k$  gives:

$$\begin{aligned} &\frac{\partial}{\partial K} \Pr(V_i \leq t, S_i \geq s_K^*)|_{K=k} \\ &= -\frac{\partial}{\partial K} \Pr(V_i \leq t, S_i \leq s_K^*)|_{K=k} = -\left[\frac{\partial}{\partial S} \Pr(V_i \leq t, S_i \leq S)|_{S=s_k^*}\right] \left(\frac{ds_K^*}{dK}|_{K=k}\right) \\ &= -F_V(t|S = s_k^*)f_S(s_k^*) \left(\frac{ds_K^*}{dK}|_{K=k}\right) = -F_V(t|S_i = s_k^*) \left(\frac{dF_S(s_k^*)}{dK}|_{K=k}\right), \end{aligned}$$

where we have used the independence between  $(V_i, S_i)$  and entry costs. [Gentry and Li (2013) used this derivative-based argument in their derivation for bounds on value distributions in auction models with endogenous entry.] Hence

$$F_V(t|S_i = s_k^*) = -\frac{\frac{\partial}{\partial K} \Pr(V_i \leq t, i \text{ enters}|K)|_{K=k}}{\frac{d}{dK} \Pr(i \text{ does not enter}|K)|_{K=k}}$$

because  $F_S(s_k^*) = \Pr(i \text{ does not enter}|k)$  in the pure-strategy BNE and  $\frac{d}{dK} F_S(s_K^*) = \frac{d}{dK} \Pr(i \text{ does not enter}|K)$ . The denominator is non-zero under the assumption of the proposition. Hence  $F_V(t|S_i = s_k^*)$  is identified for  $t \geq r$  as long as  $\Pr(W \leq t|k, a)$  and  $\Pr(i \text{ enters}|k)$  are observed in an open neighborhood around  $k$  for some  $a \geq 2$ .  $\square$

*Proof of Proposition 7.* Because  $(V_i, S_i)$  are independent across bidders, the joint distribution of  $(V_i, P_i)$  conditional on  $A_{-i} = \{j \neq i : S_j \geq s_k^*\}$  and  $S_i = s_k^*$ , evaluated at  $(P_i, V_i) = (p, v)$ , are factored as:

$$F_{P_i|A_{-i}=\{j \neq i: S_j \geq s_k^*\}}(p)F_{V_i|S_i=s_k^*}(v). \quad (\text{A13})$$

To identify  $\bar{\pi}$ , it suffices to recover the two conditional distributions in (A13) at all  $p$  and  $v \geq r$ . Because  $(V_i, S_i)$  are identically distributed across  $i$ , for any  $t \geq r$ ,

$$\begin{aligned} & \Pr(P_i \leq t | A_{-i} = \{j \neq i : S_j \geq s_k^*\}) \\ &= \sum_{m=0}^{N-1} [F_{V|S \geq s_k^*}(t)]^m \binom{N-1}{m} F_S(s_k^*)^{N-m-1} [1 - F_S(s_k^*)]^m \end{aligned} \quad (\text{A14})$$

where  $N$  is the number of potential bidders (including  $i$ ). The subscript  $i$  is suppressed to simplify notations. By construction,  $F_S(s_k^*) = \Pr(i \text{ enters} | k)$  is identified. Let  $a$  denote the realized number of entrants. First, under Assumption S,  $F_{V|S \geq s_k^*}(t)$  is over-identified for  $t \geq r$  from the distribution of prices  $W$  under  $k$  and *any*  $a \geq 2$ . (See Lemma A3 of Appendix A for details.) Next, Lemma A4 in Appendix A shows that, under Assumption S,  $F_V(t|S_i = s_k^*)$  is identified for any  $t \geq r$  provided that for some  $a \geq 2$ ,  $\Pr(W \leq t | k, a)$  and  $\Pr(i \text{ enters} | k)$  are observed for all  $k'$  in an open neighborhood around  $k$ . This is because the independence between private information and entry costs implies  $\frac{\partial}{\partial K} \Pr(V_i \leq t \text{ and } i \text{ enters} | K) |_{K=k} = -F_V(t|S = s_k^*) \left[ dF_S \left( s_k^* \right) / dK |_{K=k} \right]$ . Thus the conditional distributions in (A13), as well as  $\bar{\pi}(s_k^*)$ , are identified.  $\square$

## B Limiting Distribution of Test Statistic

First off, we show  $E[(V_i - P_i)_+ | A_{-i} = a] = \int_r^{\bar{v}} F_V(v)^a [1 - F_V(v)] dv$  under Assumption R. With a slight abuse of notation, for any generic integrable function  $\zeta$ , we use  $\int_r^{\bar{v}} \zeta(s) ds$  and  $\int_r^{\cdot} \zeta(s) ds$  as shorthands for the improper integrals  $\lim_{v \uparrow \bar{v}} \int_r^v \zeta(s) ds$  and  $\lim_{v \downarrow r} \int_v^{\cdot} \zeta(s) ds$  respectively. By definition,  $E[(V_i - P_i)_+ | A_{-i} = a]$  is

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

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  $E[(V_i - P_i)_+ | A_{-i} = a]$  as:

$$\begin{aligned} & \int_r^{\bar{v}} \left[ (\bar{v} - p)q_0 + \int_p^{\bar{v}} (v - p) dF_V(v) \right] dF_{P_i|a}(p) + F_V(r)^a \left[ \int_r^{\bar{v}} (v - r) dF_V(v) + (\bar{v} - r)q_0 \right] \\ &= q_0 \left[ \int_r^{\bar{v}} (\bar{v} - p) dF_{P_i|a}(p) \right] + \int_r^{\bar{v}} \left[ \int_p^{\bar{v}} (v - p) dF_V(v) \right] dF_{P_i|a}(p) \\ & \quad + F_V(r)^a \left[ \int_r^{\bar{v}} (v - r) dF_V(v) + (\bar{v} - r)q_0 \right]. \end{aligned} \quad (\text{B15})$$

Using integration by parts, the first term on the R.H.S. of (B15) is

$$q_0 \left[ 0 - (\bar{v} - r)F_{P_i|a}(r) + \int_r^{\bar{v}} F_{P_i|a}(p) dp \right]. \quad (\text{B16})$$

By an application of the Bounded Coverage Theorem and integration by parts, the second term on the R.H.S. of (B15) is:

$$\begin{aligned}
& \lim_{s \rightarrow r^+} \lim_{v' \rightarrow \bar{v}^-} \lim_{v'' \rightarrow \bar{v}^-} \int_s^{v'} \left[ \int_p^{v''} (v-p) dF_V(v) \right] dF_{P_i|a}(p) \\
= & \lim_{v' \rightarrow \bar{v}^-} \lim_{v'' \rightarrow \bar{v}^-} \left[ \int_{v'}^{v''} (v-v') dF_V(v) \right] F_{P_i|a}(v') - \lim_{s \rightarrow r^+} \lim_{v'' \rightarrow \bar{v}^-} \left[ \int_s^{v''} (v-s) dF_V(v) \right] F_{P_i|a}(s) \\
& + \lim_{s \rightarrow r^+} \lim_{v' \rightarrow \bar{v}^-} \lim_{v'' \rightarrow \bar{v}^-} \int_s^{v'} F_{P_i|a}(p) [F_V(v'') - F_V(p)] dp \tag{B17}
\end{aligned}$$

where the second term on the R.H.S. is 0 and 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 (B17) is 0 while the second equals  $\left[ \int_r^{\bar{v}} (v-r) dF_V(v) \right] F_{P_i|a}(r)$ . By another application of the Bounded Coverage Theorem, the third term of (B17) is:

$$\lim_{s \rightarrow r^+} \lim_{v' \rightarrow \bar{v}^-} \int_s^{v'} [F_V(p)]^a [F_V(p) - \lim_{v'' \rightarrow \bar{v}^-} F_V(v'')] dp = \int_r^{\bar{v}} [F_V(p)]^a [1 - q_0 - F_V(p)] dp. \tag{B18}$$

Substituting (B16)-(B18) into (B15), we get:

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

We now introduce some notations used for deriving the limiting distribution of the test statistic in Section 4.2. 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  $F_{W|m}$  denote the distribution of  $W$  given  $A = m$ , and  $F_{\tilde{W}|m}$  denote  $\{F_{W|m}(t) : t \in [r, \bar{v}]\}$ . Let  $\hat{F}_{\tilde{W}|m,T}$  denote estimators for  $F_{\tilde{W}|m}$  as defined in Section 4.2. 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,  $F_{W|m}(s) \equiv \Pr(W \leq s | A = m) = \pi_m^{-1} \mathbb{P}_0 1\{W \leq s, A = m\}$ . 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  $\rho_a \equiv \Pr(A_{-i} = a)$ .

Recall the distribution  $F_{W,A}$  in the data-generating process (DGP) takes the form of a mixture distribution:  $F_{W,A} = \int F_{W,A|k} dF_K$ . First off, we characterize in Lemma B5 below the joint asymptotic property of three building blocks of our test statistic:  $\hat{F}_{\tilde{W}|m,T}$  for  $2 \leq m \leq N$ ,  $\hat{\rho}_T$  and  $\frac{1}{T} \sum_{t \leq T} \tilde{K}_t$ . Introduction of some notations in the characterization are in order. For  $2 \leq m \leq N$ , let  $\{\mathbb{F}_{\tilde{W}|m} : 2 \leq m \leq N\}$  denote  $N-1$  zero-mean Gaussian Processes, each of which is indexed by  $[r, \bar{v}]$ . For any  $m \geq 2$ , covariance between  $\mathbb{F}_{\tilde{W},m}(s)$  and  $\mathbb{F}_{\tilde{W},m}(v)$  is:

$$\tilde{\Sigma}_{s,v,m} \equiv \begin{bmatrix} F_{W|m}(s) (1 - F_{W|m}(s)) & F_{W|m}(s) (1 - F_{W|m}(v)) \\ F_{W|m}(s) (1 - F_{W|m}(v)) & F_{W|m}(v) (1 - F_{W|m}(v)) \end{bmatrix}$$

for any  $s < v$  on  $[r, \bar{v})$ . Let  $[\mathcal{N}'_\rho, \mathcal{N}'_\mu]'$  denote a multivariate normal random vector in  $\mathbb{R}^{N+1}$ , where  $\mathcal{N}'_\rho$  is a  $N$ -vector with variance of the  $a$ -th coordinate (denoted  $\mathcal{N}'_{\rho,a}$ ) being  $\rho_a(1 - \rho_a)$  and covariance between the  $a$ -th and  $b$ -th coordinates being  $\frac{b(N-b)}{N^2}\pi_b - \rho_a\rho_b$  if  $b = a + 1$  and  $-\rho_a\rho_b$  for  $b \neq a + 1$ . The variance of the scalar variable  $\mathcal{N}'_\mu$  is  $Var(K) + Var(\epsilon)$ , and its covariance with  $\mathcal{N}'_{\rho,a}$  is  $E[1\{A_{-i} = a\}K] - \rho_a\mu_K$ . Covariance between the processes  $\{F_{\tilde{W}|m} : 2 \leq m \leq N\}$  and  $[\mathcal{N}'_\rho, \mathcal{N}'_\mu]'$  satisfies: (i)  $Cov(\mathbb{F}_{\tilde{W}|m}(s), \mathbb{F}_{\tilde{W}|m'}(s')) = 0$  for all  $m \neq m'$  and  $s, s' \geq r$ ; (ii)  $Cov(\mathbb{F}_{\tilde{W}|m}(s), \mathcal{N}'_{\rho,a}) = 0$  for all  $s$  and  $2 \leq m \leq N$ ,  $0 \leq a \leq N - 1$ ; (iii)  $Cov(\mathbb{F}_{\tilde{W},m}(s), \mathcal{N}'_\mu) = \frac{1}{\sqrt{\pi(m)}}(E[1\{W \leq s, A = m\}K] - F_{W|m}(s)E[1\{A = m\}K])$ ; and (iv)  $Cov(\mathcal{N}'_{\rho,a}, \mathcal{N}'_\mu) = E[1\{A_{-i} = a\}K] - \rho_a\mu_K$ .<sup>21</sup>

**Lemma B5** *Suppose  $r$  is binding and  $F_V$  is continuously distributed with positive densities over  $[r, \bar{v})$ . Suppose Assumption R holds; and  $0 < \lambda_k^* < 1$  for all  $k \in [\underline{k}, \bar{k}]$ . Then*

$$\sqrt{T} \begin{pmatrix} \hat{F}_{\tilde{W}|2,T} - F_{\tilde{W}|2} \\ \vdots \\ \hat{F}_{\tilde{W}|N,T} - F_{\tilde{W}|N} \\ \hat{\rho}_T - \rho \\ \frac{1}{T} \sum_{t \leq T} \tilde{K}_t - \mu_K \end{pmatrix} \rightsquigarrow \begin{pmatrix} \frac{1}{\sqrt{\pi_2}} \mathbb{F}_{\tilde{W},2} \\ \vdots \\ \frac{1}{\sqrt{\pi_N}} \mathbb{F}_{\tilde{W},N} \\ \mathcal{N}'_\rho \\ \mathcal{N}'_\mu \end{pmatrix}. \quad (\text{B19})$$

*Proof of Lemma B5.* Define three classes of functions with domain over the support of  $(W, A)$ :

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

Let  $\mathcal{F}_{W,A} \equiv \{f_W \wedge f_A : f_W \in \mathcal{F}_W, f_A \in \mathcal{F}_A\} \equiv \{1\{W \leq s, A = m\} : s \in [r, \bar{v}), 0 \leq m \leq N\}$  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  $\mathcal{F} \equiv \mathcal{F}_{W,A} \cup \mathcal{F}_\rho$  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  $f, f' \in \mathcal{F}$ .

Let  $\kappa$  be a mapping from  $\mathcal{B}([r, \bar{v})) \otimes (0, 1)^{N-1} \otimes (0, 1)^N$  to  $\mathcal{B}([r, \bar{v}) \otimes \{2, \dots, N-1\}) \otimes (0, 1)^N$  such that  $\kappa(F, \pi, \rho)$  evaluated at any  $s \in [r, \bar{v})$ ,  $m \in \{2, \dots, N-1\}$  and  $a \in \{0, \dots, N-1\}$  is:

$$\kappa(F, \pi, \rho)(s, m, a) \equiv \left( \frac{F(s, m)}{\pi_m}, \rho_a \right).$$

<sup>21</sup>It is worth noting that uncorrelation in (ii) holds for *all*  $a, m$  including  $m = a - 1, a, a + 1$ . This is due to the assumption that  $\epsilon$  (error in  $\tilde{K}$ ) is orthogonal to  $(W, A)$ .

With a slight abuse of notation, we also write  $\kappa(F, \pi, \rho)(s, m, a)$  as  $\kappa(F(s, m), \pi_m, \rho_a)$  below. Under the assumption of Lemma B5,  $\rho_a > 0$  for  $0 \leq a \leq N - 1$ . Thus  $\kappa$  is Hadamard-differentiable at the true  $(\{F_{W,m} : 2 \leq m \leq N\}, \pi, \rho)$  in data-generating process tangentially to the domain of  $\kappa$ , with the derivative at the truth being:

$$\kappa'(h_1, h_2, h_3)(s, m, a) = \left( \frac{1}{\pi_m} h_1(s, m) - \frac{\mathbb{P}_0 1\{W \leq s, A=m\}}{\pi_m^2} h_2(m), h_3(a) \right)$$

for  $h_1 \in \mathcal{B}([r, \bar{v}))$ ,  $h_2 \in (0, 1)^{N-1}$  and  $h_3 \in (0, 1)^N$ . An application of the Functional Delta Method (Theorem 2.8 in Kosorok (2008)) applies to characterize the weak convergence in (B19). The covariance function is derived by applying a finite-dimensional, multivariate delta method to the limiting process evaluated at finite sets of elements from the index set  $[r, \bar{v}) \otimes \{2, \dots, N\} \otimes \{0, \dots, N - 1\}$ . Mutual independence between the zero-mean Gaussian processes  $\{\mathbb{F}_{\tilde{W},m}\}_{m=2}^N$  and their joint independence from  $\mathcal{N}_\rho$  follow from an application of the delta method to random vectors. Covariance between  $\mathcal{N}_\mu$  and Gaussian processes  $\{\mathbb{F}_{\tilde{W},m}\}_{m=2}^N$  and  $\mathcal{N}_\rho$  also follows from an application of the multivariate delta method to  $\hat{F}_{\tilde{W}|m,T}(s), \hat{\rho}_{T,a}$  and sample mean of  $\tilde{K}_t$  for any  $s, m, a$ .  $\square$

The next step is to characterize the joint limiting behavior of  $\hat{F}_{V,T}$  and  $\hat{\rho}_T$ . Let  $\hat{F}_{V,m,T}(s) \equiv \phi_m^{-1}(\hat{F}_{W|m,T}(s))$ . 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  $\xi'_m(F_{W|m}(s))$  and  $\xi'_m(F_{W|m}(v))$ . 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  $\frac{\mathbf{D}_{0,m} \tilde{\mathbf{\Sigma}}_{s,v,m} \mathbf{D}'_{0,m}}{\pi_m}$ . Define:

$$\mathbf{D}_1 \equiv \begin{bmatrix} \frac{1}{N-1} & 0 & \frac{1}{N-1} & 0 & \cdots & \frac{1}{N-1} & 0 \\ 0 & \frac{1}{N-1} & 0 & \frac{1}{N-1} & \cdots & 0 & \frac{1}{N-1} \end{bmatrix}.$$

2-by-(2N-2)

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

$$\begin{bmatrix} \frac{1}{N-1} & \cdots & \frac{1}{N-1} & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} \mathbf{D}_\xi \begin{bmatrix} \frac{1}{\pi_2} & \frac{-p_{s2}}{\pi_2} & \cdots & 0 & 0 & 0 \\ & & \ddots & & & \\ 0 & 0 & \cdots & \frac{1}{\pi_N} & \frac{-p_{sN}}{\pi_N} & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix}$$

N-by-(2N-1)

where  $\mathbf{D}_\xi$  is a  $N$ -by- $N$  diagonal matrix with diagonal matrix being  $[\xi'_2(F_{W,2}(s)), \dots, \xi'_N(F_{W,N}(s)), 1]$  and  $p_{sm} \equiv \mathbb{P}_0 1\{W \leq s, A = m\}$  for  $s \geq r$



**Lemma B6** *Suppose  $F_{\bar{v}}$  is continuously distributed with positive densities over  $[r, \bar{v})$ . Under the conditions for Lemma B5,*

$$\sqrt{T} \left( \hat{F}_{\bar{v},T} - F_{\bar{v}} \right) \rightsquigarrow \mathbb{G}_V, \quad (\text{B20})$$

where  $\mathbb{G}_V$  is a zero-mean Gaussian Process indexed by  $[r, \bar{v})$  and is independent from  $\mathcal{N}_\rho$ . The covariance kernel for  $\mathbb{G}_V$  is  $(\mathbb{G}_V(s), \mathbb{G}_V(v)) = \mathbf{D}_1 \Sigma_{s,v} \mathbf{D}_1'$  for any  $s, v \in [r, \bar{v})$ ; and  $(\mathbb{G}_V(s), \mathcal{N}_\mu)$  is bivariate normal with covariance  $\mathbf{D}_2 \Sigma_s \mathbf{D}_2'$  for any  $s \in [r, \bar{v})$ .

Proof of Lemma B6 follows from the differentiability of  $\xi$  and an application of the multi-variate delta method, and is omitted here for brevity. Next, we prove Proposition 3. Recall  $\mathcal{S}_{[r, \bar{v})}$  denotes the set of positive, bounded and integrable Cadlag functions over  $[r, \bar{v})$ . Equipped with a sup-norm,  $\mathcal{S}_{[r, \bar{v})}$  is normed linear spaces with a non-degenerate interior.

*Proof of Proposition 3.* It follows from Lemma 20.10 in van der Vaart (1998) that the mapping  $\varphi$  as defined in Section 4.2 is Hadamard differentiable at  $F_{\bar{v}}$  tangentially to  $\mathcal{S}_{[r, \bar{v})}$ , with the derivative  $D_{\varphi, F_{\bar{v}}} : \mathcal{S}_{[r, \bar{v})} \rightarrow \mathbb{R}_+^N$  being

$$D_{\varphi, F_{\bar{v}}}(h)(a) \equiv \int_r^{\bar{v}} \{a[F_{\bar{v}}(s)]^{a-1}h(s) - (a+1)[F_{\bar{v}}(s)]^a h(s)\} ds$$

for any  $h \in \mathcal{S}_{[r, \bar{v})}$  and  $a \geq 1$ ; and  $D_{\varphi, F_{\bar{v}}}(h)(0) \equiv -\int_r^{\bar{v}} h(s) ds$ . The Jacobian of  $\tau$  with respect to its components  $(\varphi, \rho, \mu_K)$  at their true DGP values as:

$$[\rho_0, \dots, \rho_{N-1}, \varphi_0, \dots, \varphi_{N-1}, -1] \equiv [\rho, \varphi, -1].$$

Since  $\varphi$  is Hadamard differentiable at  $F_{\bar{v}}$  tangentially to  $\mathcal{S}_{[r, \bar{v})}$ , it follows from Lemma B6 and the Functional Delta Method (Theorem 2.8 in Kosorok (2008)) that

$$\sqrt{T} \begin{pmatrix} \hat{\varphi}_T - \varphi \\ \hat{\rho}_T - \rho \\ \frac{1}{T} \sum_t \tilde{K}_t - \mu_K \end{pmatrix} \rightsquigarrow \begin{pmatrix} D_{\varphi, F_{\bar{v}}}(\mathbb{G}_V) \\ \mathcal{N}_\rho \\ \mathcal{N}_\mu \end{pmatrix} \quad (\text{B21})$$

where  $\mathbb{G}_V$  is a zero-mean Gaussian Process defined in Lemma B6 and is independent from  $\mathcal{N}_\rho$ . The covariance between  $\mathcal{N}_\rho$  and  $\mathcal{N}_\mu$  is given in Lemma B5, while the covariance between  $\mathbb{G}_V$  and  $\mathcal{N}_\mu$  is given in Lemma B6. An application of the multivariate delta method shows under the assumptions of Lemma B6,  $\sqrt{T}(\hat{\tau}_T - \tau_0) \rightsquigarrow \mathcal{N}_\tau \equiv \rho' D_{\varphi, F_{\bar{v}}}(\mathbb{G}_V) + \varphi' \mathcal{N}_\rho - \mathcal{N}_\mu$ . 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  $D_{\varphi, F_{\bar{v}}}$  is a linear mapping defined over  $\mathcal{S}_{[r, \bar{v})}$ . It follows from Lemma 3.9.8. of van der Vaart and Wellner (1996) that  $\mathcal{N}_\tau$  is univariate normal with zero mean.  $\square$

# C Bootstrap Inference: Consistency and Asymptotic Validity

Let  $c_{1-\alpha/2}$  denote the actual  $1 - \alpha/2$  quantile of the limiting distribution of  $\sqrt{T}(\hat{\tau}_T - \tau_0)$  in the data generating process. We use the following steps to estimate  $c_{1-\alpha/2}$  using bootstrap methods.

*Step 1:* Calculate  $\hat{\tau}_T$  using the original estimation sample.

*Step 2:* Draw a bootstrap sample with size  $T$  from the original sample with replacement. Estimate  $\tau_0$  using this bootstrap sample and denote the estimate by  $\hat{\tau}_{T,1}$ .

*Step 3:* Repeat Step 2 for  $B$  times and denote the bootstrap estimates by  $\{\hat{\tau}_{T,b}\}_{b \leq B}$ . Find the  $1 - \alpha$  quantile of the empirical distribution of the bootstrap estimates  $\{\sqrt{T}|\hat{\tau}_{T,b} - \hat{\tau}_T|\}_{b \leq B}$ . Denote it by  $\hat{c}_{1-\alpha/2,T}$ .

Let  $G_T(\cdot; F)$  and  $G_\infty(\cdot; F)$  denote respectively the finite sample distribution and the limiting distribution of  $\sqrt{T}(\hat{\tau}_T - \tau_0)$  when the actual joint distribution of  $(W, A, \tilde{K})$  in data-generating process is given by a generic permissible distribution  $F$ . Let  $F_T$  and  $F_0$  denote the empirical and the true distribution of  $(W, A, \tilde{K})$  in the data-generating process respectively.

**Lemma C7** (*Bootstrap Consistency*) *Suppose Assumption R holds. Then*

$$\lim_{T \rightarrow \infty} \Pr \left[ \sup_{\tau} |G_T(\tau; F_T) - G_\infty(\tau, F_0)| \right] = 0$$

*regardless of bidders' risk attitudes.*

*Proof of Lemma C7.* By Proposition 3,  $\sqrt{T}(\hat{\tau}_T - \tau_0)$  converges weakly to a zero-mean normal distribution whose variance depends on  $F_0$ . Across auctions in data,  $(W_t, A_t, \tilde{K}_t)$  are i.i.d. draws from  $F_0$ , and by the Uniform Law of Large Numbers, the empirical distribution of  $(W_t, A_t, \tilde{K}_t)$  converges in probability to its population counterpart uniformly over the joint support. Also by Proposition 3, for any permissible joint distribution of  $(W, A, \tilde{K})$  in data-generating process, the limiting distribution of  $\sqrt{T}(\hat{\tau}_T - \tau_0)$  is continuous over  $\mathbb{R}$ . Next, we will show that for any sequence of permissible distributions  $H_T$  that converge to  $H_0$  in sup-norm as  $T \rightarrow +\infty$ , the finite sample distribution of  $\sqrt{T}(\hat{\tau}_T - \tau_0)$  under the DGP  $F_T$  (denoted by  $G_T(\cdot; H_T)$ ) converges to  $G_\infty(\cdot; H_0)$  pointwise on the real line. Note for any  $s \in \mathbb{R}$  and pair of permissible distributions  $H_1, H_2$ ,

$$|G_T(s; H_1) - G_\infty(s; H_2)| \leq |G_T(s; H_1) - G_\infty(s; H_1)| + |G_\infty(s; H_1) - G_\infty(s; H_2)|. \quad (\text{C22})$$

For any  $s$ , the first term on the right-hand side of (C22) converges to 0 as  $T \rightarrow +\infty$  by weak convergence of the test statistic under  $H_1$ . Let  $\|\cdot\|_\infty$  denote the sup norm. It follows from the continuity of the limiting distribution in Proposition 3 that, for any  $s \in \mathbb{R}$  and  $\varepsilon > 0$ , there exists  $\eta > 0$  (possibly depending on  $s, \varepsilon$ ) such that for any  $H_1, H_2$  with

$\|H_1 - H_2\|_\infty \leq \eta$ , the second term is smaller than  $\varepsilon$  for  $T$  large enough. Now consider any  $H_T$  converging to  $H_0$  in sup-norm as  $T \rightarrow +\infty$ . The observation above implies that for any  $\varepsilon > 0$ ,  $|G_T(s; H_T) - G_\infty(s; H_0)| \leq \varepsilon$  for  $T$  large enough (or equivalently for  $\|H_T - H_0\|_\infty$  small enough). The lemma then follows from Beran and Ducharme (1991) (or Theorem 2.1 in Horowitz (2000)).  $\square$

To simplify notations, we suppress dependence of  $\mathcal{N}_\tau$  on joint distribution of  $(W, A, \tilde{K})$  in DGP for the rest of the proof. Since the limiting distribution of  $\sqrt{T}(\hat{\tau}_T - \tau_0)$  is absolutely continuous with positive densities almost surely over  $\mathbb{R}$ , an immediate corollary of the preceding lemmas is that our bootstrap estimator  $\hat{c}_{1-\alpha/2,T}$  converges in probability to the actual  $1 - \alpha/2$  quantile of the limiting distribution ( $c_{1-\alpha/2}$ ). Such consistency holds regardless of bidders' risk attitudes in DGP.

*Proof of Proposition 4.* Suppose  $\tau_0 = c > 0$  in DGP. By definition,  $\Pr\left(\sqrt{T}\hat{\tau}_T \geq \hat{c}_{1-\alpha/2,T} | \tau_0 = c\right) = \Pr\left(\sqrt{T}(\hat{\tau}_T - \tau_0) - \hat{c}_{1-\alpha/2,T} \geq -\sqrt{T}\tau_0 | \tau_0 = c\right)$ . It follows from Proposition 3 that  $\sqrt{T}(\hat{\tau}_T - \tau_0)$  converges in distribution to a zero-mean, univariate normal  $\mathcal{N}_\tau$ . For any  $\varepsilon \in (0, 1)$ , let  $c_\varepsilon < +\infty$  denote the  $\varepsilon$ -quantile of  $\mathcal{N}_\tau$ . Since  $\hat{c}_{1-\alpha/2,T} \xrightarrow{P} c_{1-\alpha/2}$  and  $\tau_0 > 0$  under  $H_A$ , we have  $\lim_{T \rightarrow +\infty} \Pr(\hat{c}_{1-\alpha/2,T} < c_\varepsilon + \sqrt{T}\tau_0 | \tau_0 = c) \rightarrow 1$  for any  $\varepsilon \in (0, 1)$ . Hence for any  $c > 0$ ,

$$\begin{aligned} & \Pr\left(\sqrt{T}\hat{\tau}_T \geq \hat{c}_{1-\alpha/2,T} | \tau_0 = c\right) \geq \Pr\left(\sqrt{T}\hat{\tau}_T \geq \hat{c}_{1-\alpha/2,T} \text{ and } \hat{c}_{1-\alpha/2,T} < c_\varepsilon + \sqrt{T}\tau_0 | \tau_0 = c\right) \\ & > \Pr\left(\sqrt{T}(\hat{\tau}_T - \tau_0) \geq c_\varepsilon \text{ and } \hat{c}_{1-\alpha/2,T} < c_\varepsilon + \sqrt{T}\tau_0 | \tau_0 = c\right) \\ & \rightarrow \lim_{T \rightarrow \infty} \Pr\left(\sqrt{T}(\hat{\tau}_T - \tau_0) \geq c_\varepsilon | \tau_0 = c\right) = 1 - \varepsilon \end{aligned}$$

as  $T \rightarrow +\infty$ . This proves the consistency of our test under fixed alternatives of risk-averse bidders ( $H_A : \tau_0 = c$  with  $c > 0$ ). Symmetric arguments show

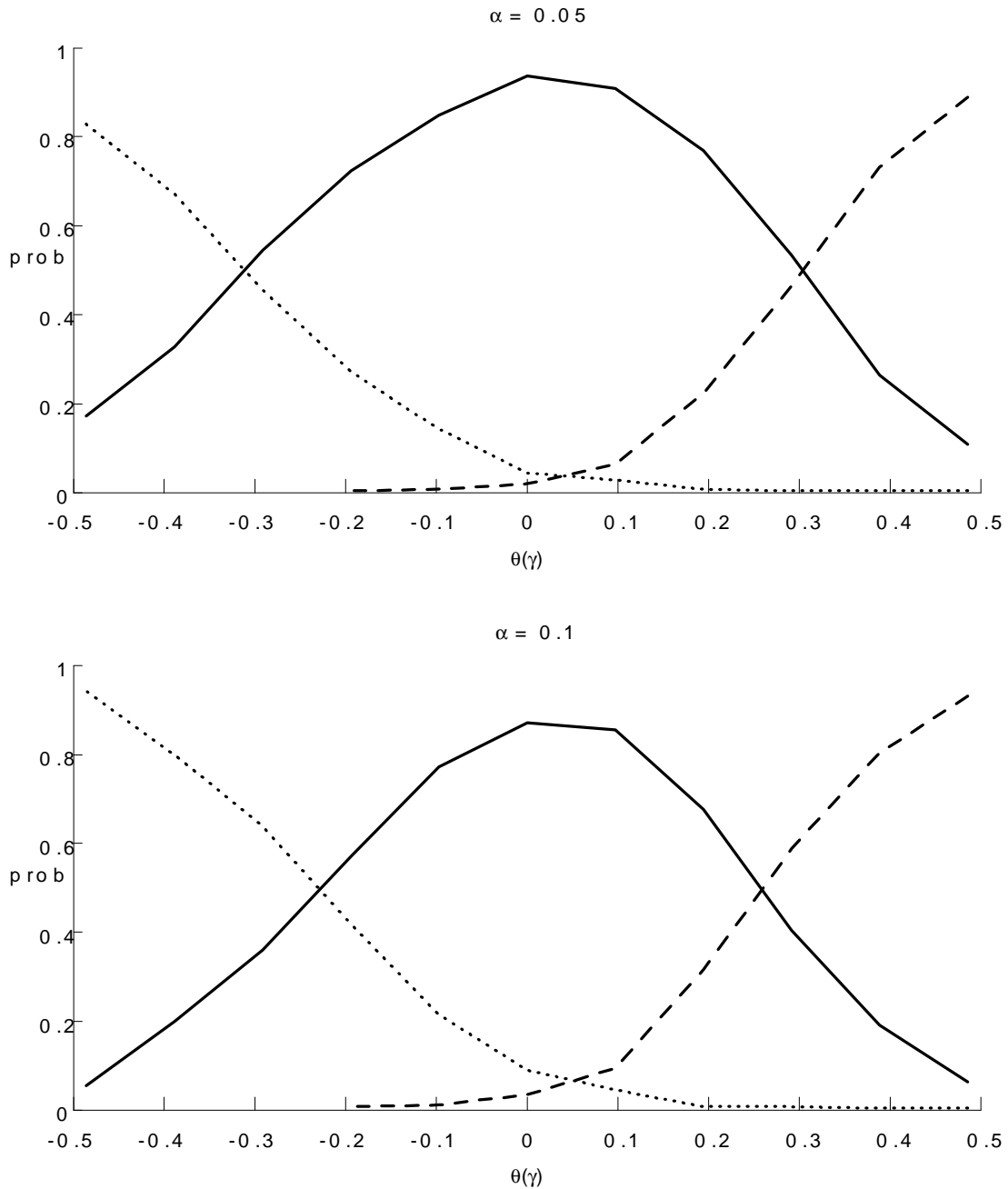
$$\lim_{T \rightarrow +\infty} \Pr\left(\sqrt{T}\hat{\tau}_T \leq -\hat{c}_{1-\alpha/2,T} | \tau_0 = c\right) = 1$$

for any  $c < 0$  (bidders are risk-loving). If bidders are risk-neutral with  $\tau_0 = 0$ ,

$$\begin{aligned} & \Pr\left(-\hat{c}_{1-\alpha/2,T} \leq \sqrt{T}\hat{\tau}_T \leq \hat{c}_{1-\alpha/2,T} | \tau_0 = 0\right) \\ & = \Pr\left(\sqrt{T}(\hat{\tau}_T - \tau_0) + \hat{c}_{1-\alpha/2,T} \geq 0 \text{ and } \sqrt{T}(\hat{\tau}_T - \tau_0) - \hat{c}_{1-\alpha/2,T} \leq 0 | \tau_0 = 0\right) \\ & \longrightarrow \Pr(-c_{1-\alpha/2} \leq \mathcal{N}_\tau \leq c_{1-\alpha/2}) = 1 - \alpha \text{ as } T \rightarrow +\infty, \end{aligned}$$

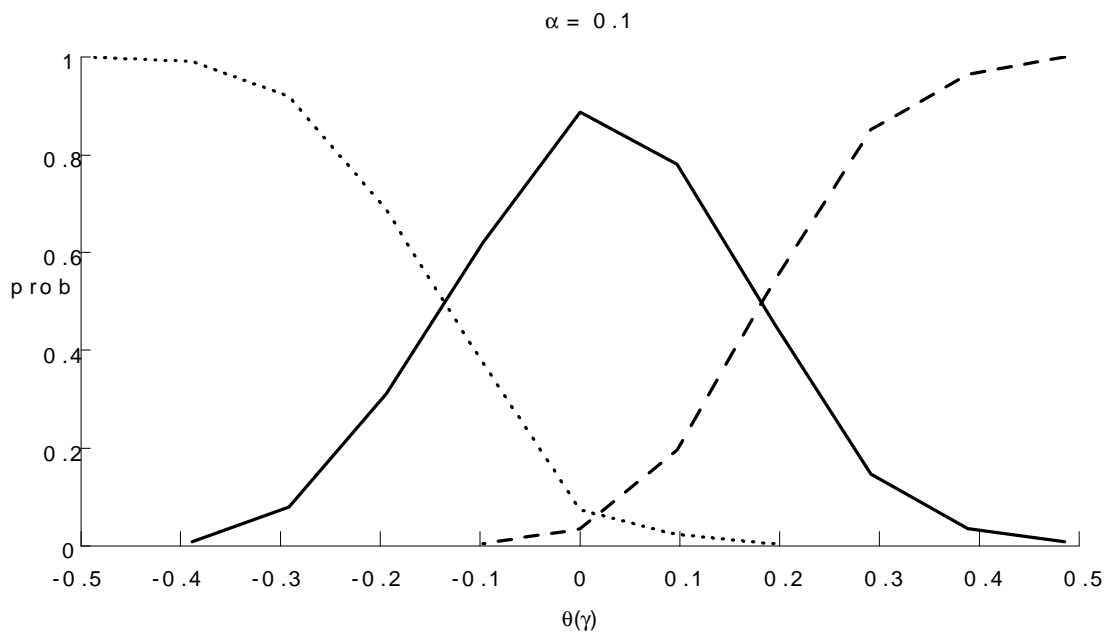
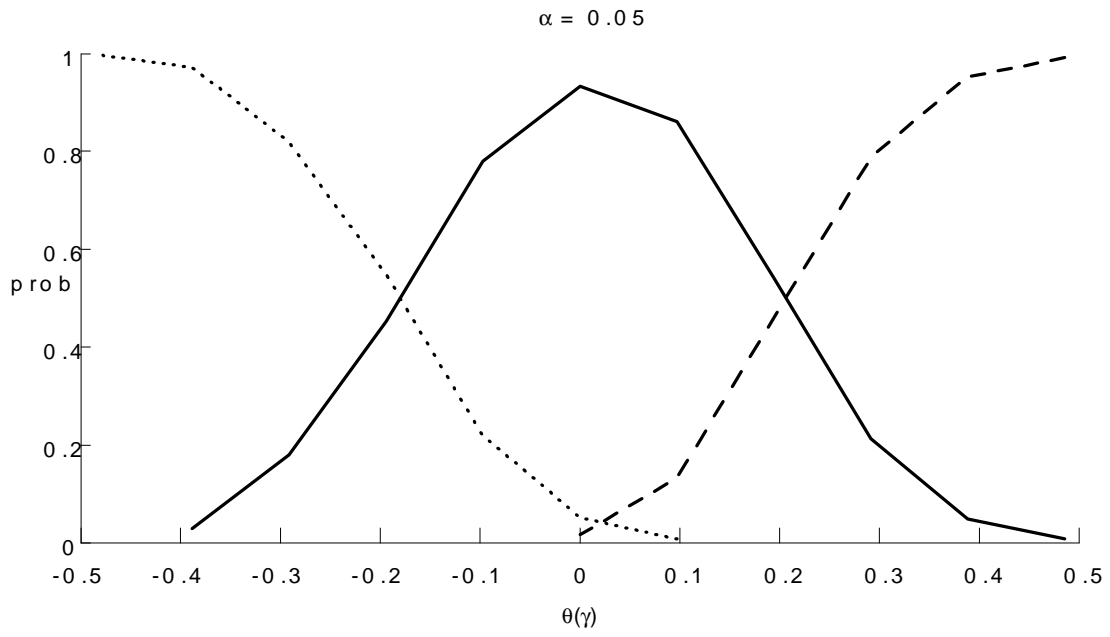
where the second equality follows from that  $\sqrt{T}(\hat{\tau}_T - \tau_0) \xrightarrow{d} \mathcal{N}_\tau$  and  $\hat{c}_{1-\alpha/2,T} \xrightarrow{P} c_{1-\alpha/2}$  ( $1 - \alpha/2$  quantile of the zero-mean normal variable  $\mathcal{N}_\tau$ ) and an application of the Continuous Mapping Theorem. *Q.E.D.*

## D Figures

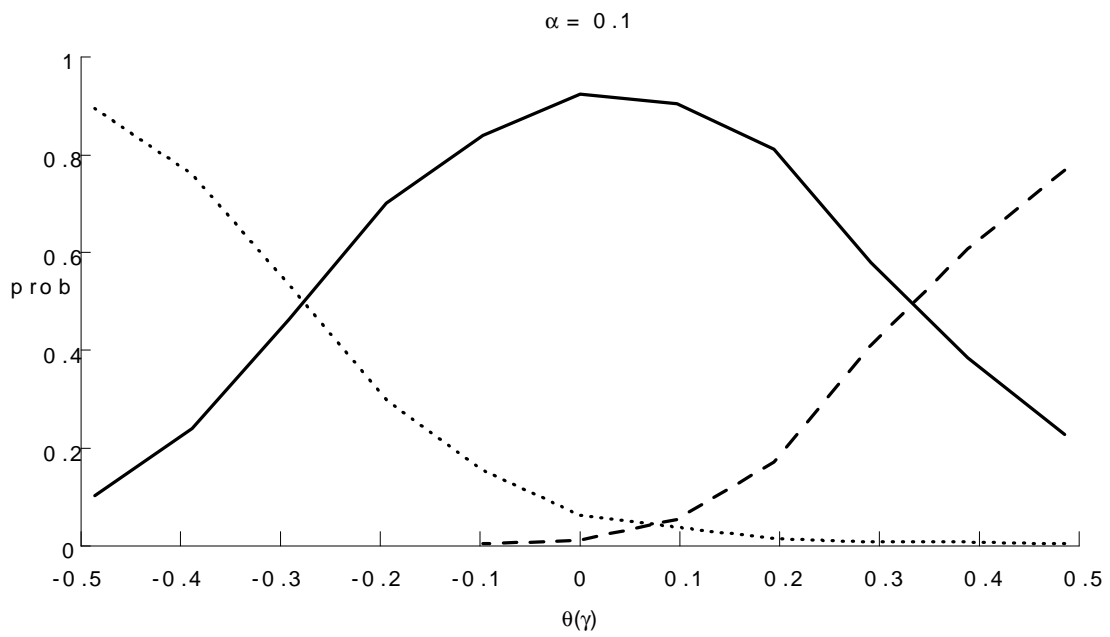
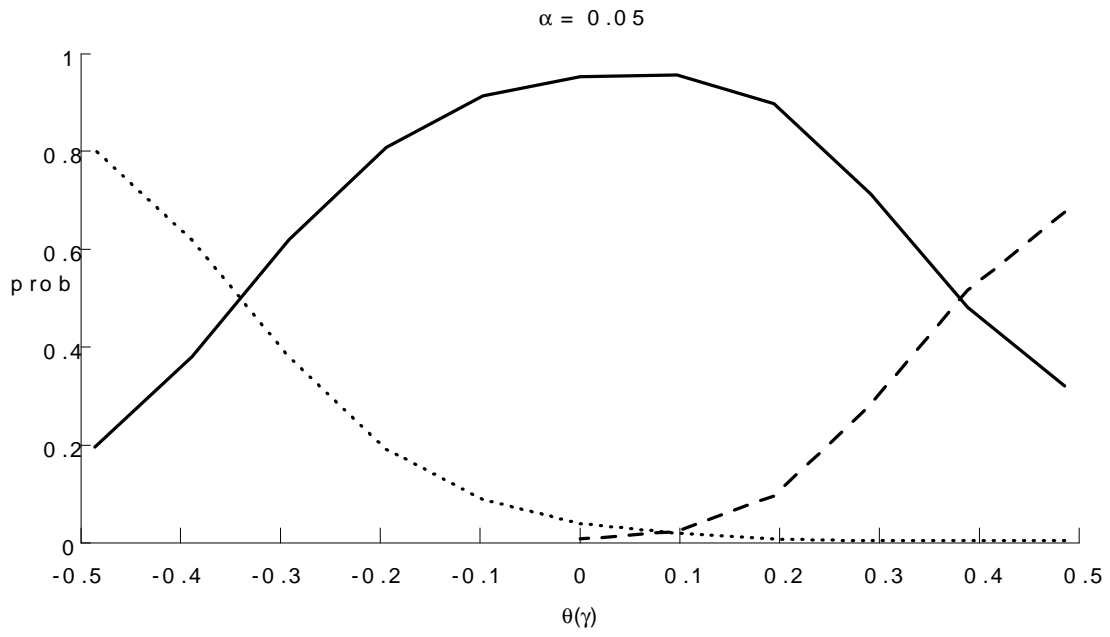


**Figure 1 (a): Test Performance** ( $N = 4, T = 1500$ )

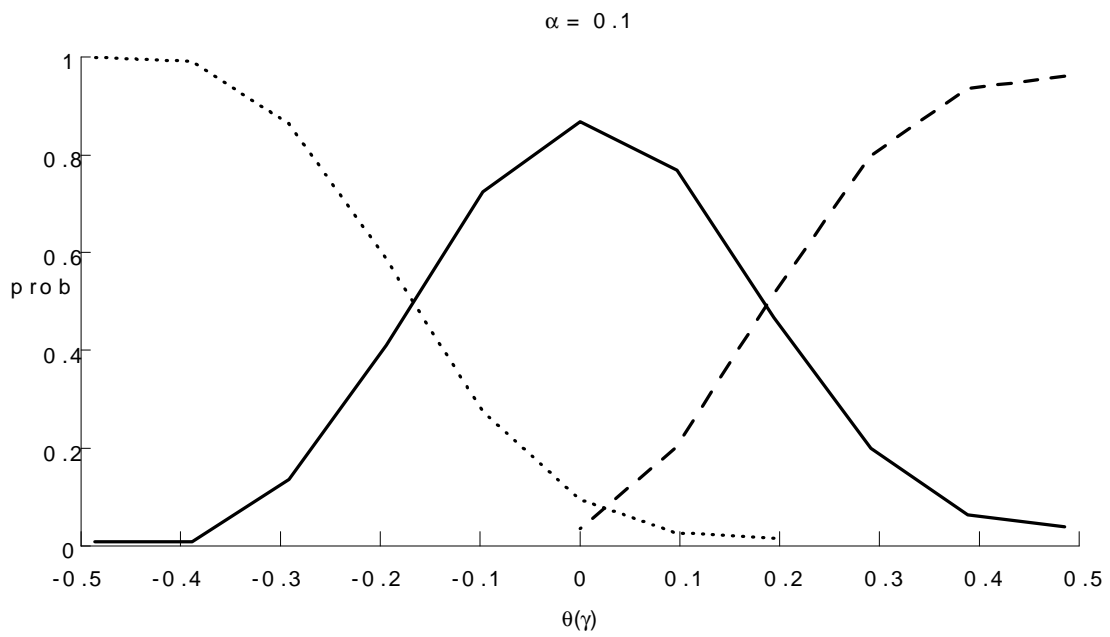
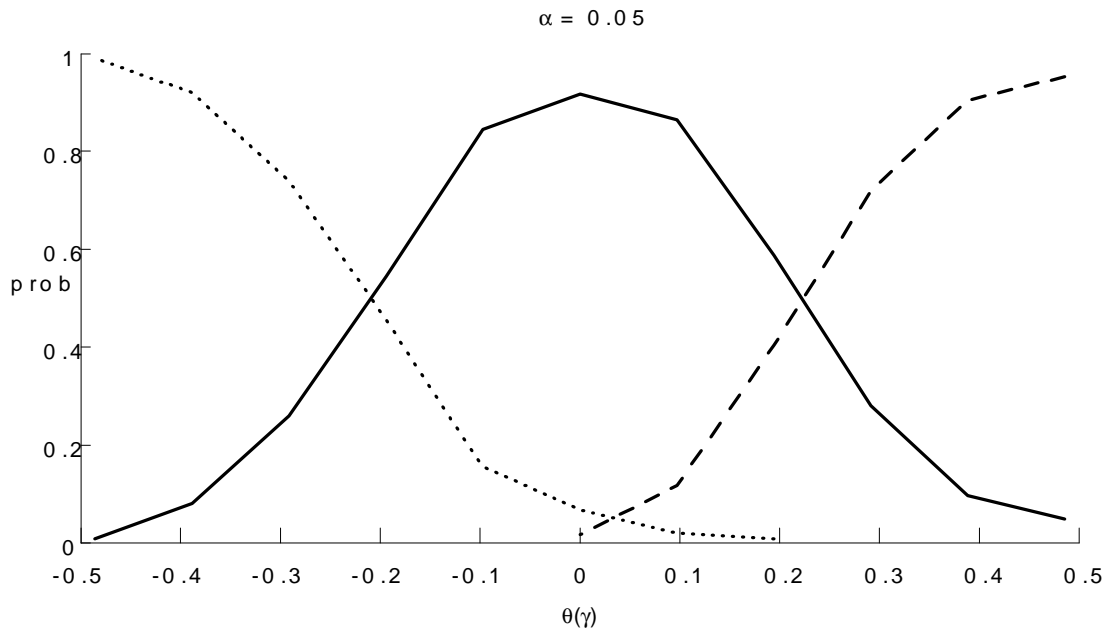
Notes: Horizontal axis: Integrated Measure of Risk Aversion. Solid line: proportion of  $S = 250$  simulated samples in which our test fails to reject the null (risk-neutrality). Dashed line: proportion that the null is rejected in favor of  $H_A$  (risk-aversion). Dotted line: proportion that the null 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 ( $N = 5, T = 3000$ )**

## References

- [1] An, Y., Hu, Y. and M. Shum (2010). "Estimating First-price Auctions with an Unknown Number of Bidders: A Misclassification Approach," *Journal of Econometrics*, 157, p328-341.
- [2] Akerberg, Daniel, Hirano, Keisuke and Quazi Shahriar (2011). "Identification of Time and Risk Preferences in Buy Price Auctions," Working Paper, University of Michigan and University of Arizona and San Diego State University
- [3] Aradillas-Lopez, Andres, Gandhi, Amit and Daniel Quint (2012). "Identification and Inference in Ascending Auctions with Correlated Private Values," *Econometrica*, forthcoming
- [4] Athey, Susan and Philip Haile (2007). "Nonparametric Approaches to Auctions," in *Handbook of Econometrics*, Vol. 6A, J. Heckman and E. Leamer, eds., Elsevier, 3847-3966.
- [5] Athey, Susan, Levin, Jonathan and E. Seira (2011). "Comparing Open and Sealed Bid Auctions: Theory and Evidence from Timber Auctions," *Quarterly Journal of Economics*, 126(1), 207–257.
- [6] Bajari, Patrick and Ali Hortacsu (2005). "Are Structural Estimates of Auction Models Reasonable? Evidence from Experimental Data," *Journal of Political Economy*, 113(4), 703-741.
- [7] Beran, R and G R Ducharme (1991). *Asymptotic Theory for Bootstrap Methods in Statistics* (Les Publications CRM, Centre de recherches mathématiques, Université de Montréal, Montréal, Canada).
- [8] Campo, Sandra, Guerre, Emmanuel, Perrigne, Isabelle and Quang Vuong (2011). "Semiparametric Estimation of First-Price Auctions with Risk Averse Bidders," *Review of Economic Studies*, Vol 78, 112-147
- [9] Carrasco, M., Florens, J.-P., Renault, E. (2006). "Linear Inverse Problems in Structural Econometrics Estimation Based on Spectral Decomposition and Regularization". In: Heckman, J.J., Leamer, E.E. (Eds.), *Handbook of Econometrics*, vol.6. North-Holland, Amsterdam
- [10] Gentry, Matthew and Tong Li (2013). "Identification in Auctions with Selective Entry," forthcoming, *Econometrica*



- [11] Guerre, Emmanuel, Perrigne, Isabelle, and Quang Vuong (2009). "Nonparametric Identification of Risk Aversion in First-Price Auctions Under Exclusion Restrictions," *Econometrica*, Vol. 77, No.4, 1193-1227.
- [12] Haile, Philip, Hong, Han and Matthew Shum (2004). "Nonparametric Tests for Common Values in First-Price Sealed-Bid Auctions." Working Paper, Yale University.
- [13] Philip A. Haile and Elie Tamer (2003). "Inference with an Incomplete Model of English Auctions," *Journal of Political Economy*, Vol. 111, No.1, 1-51
- [14] Hendricks, Kenneth, Pinkse, Joris and Robert H. Porter (2003). "Empirical Implications of Equilibrium Bidding in First-Price, Symmetric, Common-Value Auctions." *Review of Economic Studies*, 70: 115–146.
- [15] Hendricks, Kenneth and Robert Porter (2007). "A Survey of Empirical Work in Auctions," in *Handbook of Industrial Organization*, Vol. III, edited by R. Porter and M. Armstrong, Vickers, Amsterdam: North- Holland.
- [16] Hong, Han and Matthew Shum (2002) "Increasing Competition and the Winner's Curse: Evidence from Procurement," *Review of Economic Studies*, 69(4), 871–898.
- [17] Joel L. Horowitz, Chapter 52 The Bootstrap, In: J.J. Heckman and E. Leamer, Editor(s), *Handbook of Econometrics*, Elsevier, 2001, Volume 5, Pages 3159-3228
- [18] Jun, Sung Jae, Pinkse, Joris and Yuanyuan Wan (2010). "A consistent nonparametric test of affiliation in auction models," *Journal of Econometrics* 159 (2010), 46-54.
- [19] Kosorok, M. (2008). *Introduction to Empirical Processes and Semiparametric Inference*. Springer-Verlag: New York.
- [20] Krasnokutskaya, Elena and Katja Seim, "Bid Preference Programs and Participation in Highway Procurement," *American Economic Review*, 101, 2011
- [21] Levin, Dan and James L. Smith (1994). "Equilibrium in Auctions with Entry," *American Economic Review*, Vol. 84, No. 3, 585-599.
- [22] Tong Li (2005). "Econometrics of First-Price Auctions with Entry and Binding Reservation Prices," *Journal of Econometrics*, 126, 173-200
- [23] Li, Tong, Lu, Jingfeng and Li Zhao (2012). "Auctions with Selective Entry and Risk Averse Bidders: Theory and Evidence," Working Paper, Vanderbilt University and National University of Singapore

- [24] Li, Tong and X. Zheng (2009). "Entry and Competition Effects in First-Price Auctions: Theory and Evidence from Procurement Auctions," *Review of Economic Studies*, 76, 1397-1429.
- [25] Li, Tong and B. Zhang (2010). "Testing for Affiliation in First-Price Auctions Using Entry Behavior," *International Economic Review*, 51, 837-850.
- [26] Lu, Jingfeng and Isabelle Perrigne (2008). "Estimating Risk Aversion from Ascending and Sealed-Bid Auctions: The Case of Timber Auction Data," *Journal of Applied Econometrics*, 23, 871-896.
- [27] Marmer, Vadim, Artyom Shneyerov, and Pai Xu (2013). "What Model for Entry in First-Price Auctions? A Nonparametric Approach," forthcoming, *Journal of Econometrics*
- [28] Matthews, Steven (1987). "Comparing Auctions for Risk Averse Buyers: A Buyer's Point of View," *Econometrica*, Vol. 55, 633-646.
- [29] Roberts, James and Andrew Sweeting (2010). "Entry and Selection in Auctions," NBER Working Paper No. 16650.
- [30] Smith, James and Dan Levin (1996). "Ranking Auctions with Risk Averse Bidders," *Journal of Economic Theory*, Vol. 68, No. 2, 549-561.
- [31] van der Vaart, A.W., (1998). *Asymptotic Statistics*, 1-496. Cambridge University Press, Cambridge.
- [32] Van der Vaart, A. and J. Wellner (1996). *Weak Convergence and Empirical Processes*. Springer-Verlag: New York.
- [33] Lixin Ye (2007). "Indicative Bidding and A Theory of Two-Stage Auctions," *Games and Economic Behavior*, January 2007, 58 (1): 181-207.