Nonparametric inference in asymmetric first-price auctions with k-rationalizable beliefs

Lidia Kosenkova

November 14, 2019

Abstract

This paper studies bidding behavior in first-price sealed bid auctions with risk-neutral bidders and independent private values. Instead of assuming that bids and beliefs are consistent with a Bayesian Nash equilibrium (BNE), I only assume consistency with k steps of iterated elimination of strategies that are never best responses (k-rationalizability). I provide an econometric test for the largest value of k that is consistent with the data. Rejecting any finite k automatically rules out BNE and (full) rationalizability. It allows to quantify strategic sophistication of bidders and improve counterfactual predictions. My framework includes "cognitive hierarchy" or "level-k" models as special cases but, unlike those models, I make no assumptions about how beliefs are selected. My approach relies only on inequalities between functionals of conditional distributions that are implied by k-rationalizability. As an empirical illustration I apply my test to USFS timber auction data. The results show that values of k as low as 2 can be rejected in asymmetric auctions. Counterfactual exercises allow me to quantify the loss in expected payoff derived from the presence of incorrect beliefs.

1 Introduction

First-price sealed-bid auctions are widely used to sell objects ranging from drilling and logging rights, to contracts for constructing roads, etc. By the nature of these auctions, optimal bidding strategies –and their observable implications– depend crucially on the assumptions made on bidders' beliefs about opponents' behavior. By far, most of the identification results in the literature have been obtained under the assumption of *Bayesian Nash equilibrium* (BNE), which presupposes correct beliefs for bidders. As a result, identification results have been derived under BNE and they have enabled estimation and inference in first-price auction models under a variety of assumptions regarding the distribution of bidders' values (Paarsch (1992), Laffont et al. (1995), Donald and Paarsch (1996), Guerre et al. (2000), Li et al. (2000), Li et al. (2002), Hubbard et al. (2012)). However, having methods to identify and test the true underlying behavioral model is a crucial first-step to obtaining credible results. While the experimental literature has developed alternative models to BNE, these typically rely on very precise assumptions about how different types of behavior deviate from BNE.

This paper is motivated by the need for robust econometric methods that can identify and test bidding behavior in auctions without making specific assumptions about how bidders may deviate from BNE. My goal is to frame the econometric problem within a general class of behavioral models that nests BNE and rationalizability (arguably, the most important solution concepts in the literature) as special cases. To this end, while I maintain the assumption that bidders are expected-utility maximizers, I allow for incorrect beliefs and assume only that these are consistent with certain rationality requirements but are otherwise unrestricted. Specifically, following theoretical results from Battigalli and Siniscalchi (2003), I assume that beliefs are consistent with k iterated steps of elimination of bidding strategies that are never best-responses. Otherwise, beliefs are left unrestricted.

This behavioral model is called *k-rationalizability* and it includes BNE and rationalizability (common knowledge of expected-profit maximazing behavior) as special cases. Importantly, it also includes cognitive hierarchy models (often called "Level-k" models), developed in the experimental literature as alternatives to BNE, as special cases (Crawford and Iriberri (2007), Gillen (2009), Crawford et al. (2013), Kagel and Levin (2014), An (2017), Hortacsu et al. (2017)). This paper develops econometric techniques that identify and test where the underlying behavior in the data falls within the range of k-rationalizable bidding. Specifically, I identify the largest value of k that is consistent with the observed bidding behavior by testing an inequality involving transformations of conditional distributions. As a byproduct, such a procedure tests, e.g, whether the data is consistent with BNE (rejecting any finite k rejects BNE) but, more generally, it allows me to identify bidders' degree of *strategic sophistication* as measured by k. This, in turn, can be used to undertake credible counterfactual analysis.

All existing nonparametric identification results in first-price auctions are immediately invalidated in the very general behavioral setting considered here. However, my goal is not to identify the distribution of values from the observed first-price auction data (an infeasible task given my behavioral assumptions), but rather to infer where the observed bidding behavior falls within the class of k-rationalizable bidding. In this paper I show that if the distribution of values is nonparametrically identified from an outside source, a test for the largest value of k consistent with k-rationalizability in the first-price auction data can be conducted nonparametrically. To this end, I develope an econometric procedure focused on the case where the researcher observes auctions for the same objects from first-price as well as ascending auctions, since bidding one's own valuation remains a weakly dominant strategy in the latter regardless of higher-order beliefs and, consequently, existing identification results for value distributions in ascending auctions (Athey and Haile (2002), Haile and Tamer (2003)) remain valid and are not affected by k-rationalizable behavior in first-price auctions. While combining data from ascending and first-price auctions of the same type of object has been done before, e.g, in Lu and Perrigne (2008) (to estimate risk aversion) and in Athey et al. (2011) (to study bidders participation and auction design). To my knowledge this is the first paper that proposes a nonparametric test for bidding behavior and rationality.

The paper proceeds as follows. Section 2 describes the model of k-rationalizable bidding. Section 3 develops an nonparametric test for k-rationalizability that assumes the existence of auxiliary data from ascending auctions. Section 4 discusses the case where no auxiliary data is available. An empirical illustration for USFS timber auction data is included in Section 5. Classifying firms into small and large according to the number of workers, my test shows that values of k as low as 2 are rejected in two-bidder auctions with asymmetric types, while larger values of k are consistent with the data in symmetric auctions. Counterfactual analysis then quantifies the economic cost of incorrect beliefs. Section 6 summarizes results from several extensions of my main assumptions, including correlation in values, risk aversion and the case of "ambiguity", where the distribution of values is unknown to bidders. Section 7 concludes. Econometric proofs and other supplemental materials are in the Appendix.

2 A Model of k-rationalizable Bidding

My goal is to develop a method for inference in first-price auctions with independent private values (IPV) under the assumption that bidders are profit maximizers and risk-neutral. Nash equilibrium behavior in auctions assumes that players hold *correct* beliefs about the strategies of other players. I relax this assumption: I require that beliefs are *strategically sophisticated* (in a way that is described below) but not necessarily correct. Furthermore, I allow unobserved heterogeneity in bidders' beliefs and I allow for beliefs to be (possibly) dependent on bidders' values. I consider a weaker solution concept than BNE, but which includes BNE as a special case. At the same time, the notion of strategically sophisticated beliefs implies that beliefs must be justified by a well-founded rationalizability criterion. To achieve these goals, I use the results of Battigalli and Siniscalchi (2003) (henceforth B-S) about rationalizable bidding in first-price auctions. B-S impose a minimal requirement on the space of allowable beliefs. The characterization of rationalizable bids is based on the iterative deletion of bids that cannot be justified by beliefs consistent with progressively higher degrees of strategic sophistication. Rationalizable bids are those that survive arbitrarily many steps of iterative deletion, but the B-S framework also characterizes bids that survive only finitely many steps of iteration, giving rise to a well–defined notion of k-rationalizable bids.

Under conditions I describe below, the space of k-rationalizable bids are completely described by a (sharp) upper bound $\overline{B}_k(\cdot)$ on bidding functions (mappings from bidders' values to bids). A bidding function $b(\cdot)$ is consistent with k-rationalizability if and only if $\forall\ v\ b(v) \leq \overline{B}_k(v)$. These bounds turn out to have useful properties: for each k the upper bound $\overline{B}_k(\cdot)$ is a strictly increasing function, bounds are monotonically decreasing in k, so that $\overline{B}_k(\cdot) \geq \overline{B}_{k+1}(\cdot)$ for all k, and if I define $\overline{B}_\infty(v) = \lim_{k\to\infty} \overline{B}_k(v)$, then $\overline{B}_\infty(\cdot)$ is a sharp upper bound for the space of rationalizable bids. Since every BNE is rationalizable, each BNE bidding function is below the bound $\overline{B}_\infty(\cdot)$ and there is a gap between them. The lower bound for k-rationalizable bids is zero for any k. Thus, all bids below BNE and *some* bids above BNE are $k-$rationalizable.

Based on these properties, I provide econometric tools to test whether observed bidding behavior is consistent with $k-$rationalizability. This has important implications since rejecting *any* finite k would immediately reject BNE bidding behavior and (full) rationalizability.

2.1 Rationalizable bids

All results presented in this section come from B-S.

2.1.1 Basic setup

Consider a single-object first-price auction. There are n risk-neutral players with independent private values.[1] Private values are drawn from a commonly known distribution[2] F_0, with compact

[1]The analysis in B-S allows for common values under affiliation and symmetry restrictions. I focus on IPV because the nonparametric inferential method I propose in Section 3.2 relies on the IPV assumption. The parametric approach outlined in Section 6.3 can allow for common values under affiliation.

[2]I extend the results to a setting with asymmetric bidders in Section 2.3.

support $[\underline{v}, \overline{v}]$, where $0 \leq \underline{v} < \overline{v}$. Like B-S, I consider a setting where there is no binding reserve price, so I set it to zero. For simplicity I also normalize $\underline{v} = 0$. The following assumption[3] is maintained.

Assumption 2.1. The cdf F_0 is differentiable, with continuous density f_0 bounded away from zero over the support $[0, \overline{v}]$. F_0 is common knowledge.

Each bidder i observes v_i, her value of the good, and chooses a bid $b \geq 0$. The object is assigned to the bidder with the highest bid, ties are broken at random. The winner pays her bid and losers do not pay anything.

2.1.2 Beliefs and best responses

Bidders treat competitors' bids as random variables. More precisely, a particular conjecture of bidder i about the bidding behavior of player j is viewed as a function $b_j : [0, \overline{v}] \to \mathbb{R}_+$. Let $\mathcal{B}$ denote the set of all positive bounded functions with domain $[0, \overline{v}]$. The set of possible conjectures for bidder i about her competitors is $\mathcal{B}_{-i} = \prod_{j \neq i} \mathcal{B}_j$, where each $\mathcal{B}_j \subseteq \mathcal{B}$. A *belief* of player i is a probability measure μ_i on $\Delta(\mathcal{B}_{-i})$. I focus on beliefs that assign probability zero to ties. The expected payoff[4] of bidding b conditional on the private value v_i and a given belief $\mu_i \in \Delta(\mathcal{B}_{-i})$ is

$$\pi(b, v_i; \mu_i) \equiv (v_i - b) \int_{\mathcal{B}_{-i}} \mathbf{P}[b_{-i} \leq b] \, \mu_i(db_{-i}) \equiv (v_i - b)\mathbf{P}[b_{-i} \leq b | \mu_i]. \tag{2.1}$$

Let $\pi^*(v_i; \mu_i) \equiv \sup_{b \geq 0} \pi(b, v_i; \mu_i)$. If bid b is such that $\pi(b, v_i; \mu_i) = \pi^*(v_i; \mu_i)$, it is called a *best response* to the belief μ_i.

A restriction on the space of beliefs

As in B-S, I rule out cases where bidders submit completely non-informative or trivial bids because they are certain they will not win the good. Therefore, I restrict attention to the space of beliefs where every player assumes that any positive bid yields a positive probability of winning the good.

Assumption 2.2. Let

$$\Delta^+(\mathcal{B}_{-i}) = \left\{ \mu \in \Delta(\mathcal{B}_{-i}) \colon \forall \, b > 0, \int_{\mathcal{B}_{-i}} \mathbf{P}_{v_{-i}} \left[b_{-i}(v_{-i}) \leq b | v_i \right] \mu(db_{-i}) > 0 \; \forall \; v_i \in [0, \overline{v}] \right\}.$$

Then beliefs for each player i belong to $\Delta^+(\mathcal{B}_{-i})$.

[3]In the econometric inference section I impose further smoothness restrictions on F_0.

[4]I focus throughout on *interim* rationalizability, that is, bids are rationalizable for i conditional on observing v_i and a belief μ_i can depend on the private value v_i.

Focusing attention to $\Delta^+(\mathcal{B}_{-i})$ rules out cases where bidder i submits a bid above her private value v_i or a bid equal to zero simply because she is certain she will not win the good. Assuming that bidders are expected profit maximizers and (2.2) rules out weakly dominated bids[5]. This also yields a natural upper bound on rationalizable bids which will be the starting point of the iterative method described below.

2.1.3 Rationalizable bids for a given upper bound on opponents' bidding functions

Let $\overline{B} : [0, \overline{v}] \to \mathbb{R}_+$ be a nondecreasing function such that $\overline{B}(v) > 0$ for all $v \neq 0$. Suppose bidder i believes that $\overline{B}$ is an upper bound for the opponents' bids, i.e., $b_j(v) \leq \overline{B}(v)$ for all $v \in [0, \overline{v}]$ and $j \neq i$. Denote $\overline{B}_{-i} = \left\{\overline{B}, \overline{B}, \ldots, \overline{B}\right\}$. Accordingly, let

$$\Delta^+(\mathcal{B}_{-i}; \overline{B}_{-i}) = \left\{\mu \in \Delta^+(\mathcal{B}_{-i})\colon\ \mu\left(\left\{b_{-i}\colon\ b_{-i} < \overline{B}_{-i}\right\}\right) = 1\right\}.$$

Which bids can be rationalized as best responses to some beliefs in $\Delta^+(\mathcal{B}_{-i}; \overline{B}_{-i})$? For a given bid b^*, compare $v_i - b^*$ (the highest possible value of bidder i's expected payoff for bid b^* when he wins with probability 1) against $\inf_{\mu \in \Delta^+(\mathcal{B}_{-i};\overline{B}_{-i})} \pi^*(v_i; \mu)$ (the worst possible expected payoff for beliefs in $\Delta^+(\mathcal{B}_{-i}; \overline{B}_{-i})$). From Theorem 6 in B-S it follows that:

1. If $v_i - b^* < \inf_{\mu \in \Delta^+(\mathcal{B}_{-i};\overline{B}_{-i})} \pi^*(v_i; \mu)$, then b^* is not a best response to any belief $\mu \in \Delta^+(\mathcal{B}_{-i}, \overline{B}_{-i})$ given v_i.

2. If $v_i - b^* > \inf_{\mu \in \Delta^+(\mathcal{B}_{-i};\overline{B}_{-i})} \pi^*(v_i; \mu)$, then b^* is a strict best response to *some* belief $\mu \in \Delta^+(\mathcal{B}_{-i}, \overline{B}_{-i})$ given v_i.

3. The upper bound $\overline{B}$ produces a least *upper bound* on the best response function of bidder i. For each v_i, this upper bound is given by $v_i - \inf_{\mu \in \Delta^+(\mathcal{B}_{-i};\overline{B}_{-i})} \pi^*(v_i; \mu)$.

4. If $\overline{B}$ is increasing, then $\inf_{\mu \in \Delta^+(\mathcal{B}_{-i};\overline{B}_{-i})} \pi^*(v_i; \mu) = \pi^*(v_i; \overline{B}_{-i})$, where $\pi^*(v_i; \overline{B}_{-i})$ represents the optimal expected payoff for player i if her opponents' bidding functions correspond to the upper bound $\overline{B}_{-i}$.

2.1.4 An iterative characterization of rationalizable bids

The previous section presents results for a given upper bound on opponents' bids. Thus, one might wonder whether a description of rationalizable bids requires one to specify a nontrivial "initial" upper bound $\overline{B}$. Given minimal assumptions (2.1) and (2.2), this is not the case: starting with arbitrarily large upper bounds $\overline{B}$, one can arrive at the set of rationalizable bids through an iterative process of deletion of bids that are never best response.

[5] In the second-price auctions it implies that players bid their private values.

For bidder i and a given set of beliefs $\Delta_i \subseteq \Delta^+(\mathcal{B}_{-i})$, let

$$\rho_i(v_i, \Delta_i) = \{b \geq 0\colon \exists \mu \in \Delta_i, \pi(b, v_i; \mu) = \pi^*(v_i; \mu)\},$$

denote the set of bids that can be rationalized as best-responses for bidder i with private value v_i to some beliefs in Δ_i. Next, for a collection of set-valued functions $\{\mathcal{C}_j\colon [0, \overline{v}] \rightrightarrows \mathbb{R}_+\}_{j \neq i}$, let denote the set of beliefs allowed by the assumption (2.2) with the support in $\mathcal{C}_{-i} \equiv \prod_{j \neq i} \mathcal{C}_j$ by

$$\Delta^+(\mathcal{C}_{-i}) = \left\{\mu \in \Delta^+(\mathcal{B}_{-i})\colon \mu\left(\{b_j(v_j) \in \mathcal{C}_j(v_j),\ \forall\, v_j \in [0, \overline{v}],\ \forall\, j \neq i\}\right) = 1\right\}.$$

Definition 2.1. [**k-rationalizable and rationalizable**[6] **bids**] For each $i = 1, \ldots, n$ and each $v_i \in [0, \overline{v}]$, let:

$$\mathcal{R}_{i,0}(v_i) = \mathbb{R}_+, \quad \mathcal{R}_{i,k}(v_i) = \rho_i\left(v_i, \Delta^+(\mathcal{R}_{-i,k-1})\right), \quad k = 1, 2, \ldots$$

1. A bid b^* is $k-$rationalizable for bidder i given v_i if $b^* \in \mathcal{R}_{i,k}(v_i)$.
2. A bid b^* is rationalizable for bidder i given v_i if there exists an $n-$tuple $(\mathcal{C}_1, \ldots, \mathcal{C}_n)$ such that $\mathcal{C}_j(v_j) \subseteq \rho_j(v_j, \Delta(\mathcal{C}_{-j}))$ for each $j = 1, \ldots, n$, and $b^* \in \mathcal{C}_i(v_i)$.

Under assumptions (2.1), (2.2), rationalizable bids for each v_i can be obtained from the set of $k-$rationalizable bids by letting $k \to \infty$.

Upper bounds for $k-$rationalizable bids

Given the assumptions (2.1), (2.2), a key property of the set of $k-$rationalizable bids is that this set is *completely characterized by a corresponding upper bound* $\overline{B}_k(\cdot)$. For every $v_i \in [0, \overline{v}]$, let

$$\overline{B}_1(v_i) = v_i, \quad \overline{B}_{k+1}(v_i) = v_i - \inf_{\mu \in \Delta^+(\mathcal{R}_{-i,k}(\cdot))} \pi^*(v_i; \mu), \quad k = 1, 2, \ldots$$

Theorem 12 in B-S shows that the upper bound $\overline{B}_k(\cdot)$ is strictly increasing, continuous and positive for every $k \geq 1$. This increasing property in turn implies that

$$\inf_{\mu \in \Delta^+(\mathcal{R}_{-i,k}(\cdot))} \pi^*(v_i; \mu) = \pi^*(v_i; \overline{B}_k(\cdot)) = \mathop{sup}_{b \geq 0} \left\{(v_i - b)\mathbf{P}\left[\overline{B}_k(v_j) \leq b\ \forall\ j \neq i\right]\right\}$$

[6] Rationalizability is defined in terms of best-response properties (see Bernheim (1984), Pearce (1984), Fudenberg and Tirole (1998) (Section 2.3.1)) independently of the sets of $k-$rationalizable bids.

And one can re-express these bounds simply as

$$\overline{B}_1(v_i) = v_i, \quad \overline{B}_{k+1}(v_i) = v_i - \pi^*(v_i; \overline{B}_k(\cdot)), \quad k = 1, 2, \ldots. \tag{2.2}$$

Result 2.1. ***(Properties of the upper bounds for $k-$rationalizable bids)*** *From the results in Theorem 12 and Proposition 13 in B-S I have the following:*

1. $\overline{B}_k(\cdot) \geq \overline{B}_{k+1}(\cdot)$ *for all* k.

2. *For all* $k \geq 1$ *and* $v_i \in (0, \overline{v}]$, *the set of* $k-$*rationalizable bids* $\mathcal{R}_{i,k}(v_i)$ *is an interval with interior* $(0, \overline{B}_k(v_i))$. *The upper bound* $\overline{B}_k(\cdot)$ *is strictly increasing, continuous, concave and positive.*

3. *For all* $v_i \in (0, \overline{v}]$, *the set of rationalizable bids is an interval with interior* $(0, \overline{B}_\infty(v_i))$; *the upper bound* $\overline{B}_\infty(\cdot)$ *is continuous, concave, nondecreasing and positive.*

2.2 Behavior and stylized facts consistent with k-rationalizable bidding

The range of possible behavioral models encompassed within $k-$rationalizability is significant. Basically, the only initial restriction placed on beliefs is Assumption 2.2, i.e. bidders assume that any nonzero bid may win with positive probability. Thus, $k-$rationalizability includes Bayesian Nash equilibrium (symmetric or asymmetric) as a special case. It is also consistent with "level-k" or "cognitive hierarchy" models which have been used in experimental economics to explain deviations from equilibrium behavior[7] (see Crawford and Iriberri (2007), Crawford et al. (2013), Kline (2015)). Rationalizable bidding defined via best-response properties (Bernheim (1984), Osborne and Rubinstein (1994)) is also a special case of $k-$rationalizable bidding and in fact it is the limiting case as $k \to \infty$.

As B-S point out, $k-$rationalizability is also compatible with the following stylized facts that have been observed in experimental auctions:

Overbidding relative to risk-neutral BNE: The limiting bounds $\overline{B}_\infty(\cdot)$ include bids above the risk-neutral BNE bids.

Underbidding relative to risk-neutral BNE: Any bid in the interval $(0, \overline{B}_k(v_i))$ is $k-$ rationalizable. This includes bids strictly below BNE. While overbidding above risk-neutral BNE can be explained by risk-aversion (strictly concave payoff functions), bids below risk-neutral BNE cannot be originated by risk-aversion (see Section 4.1 in Krishna (2010)).

[7] If the anchor beliefs of the $L0$ type are consistent with Assumption 2.2 then $k-$rationalizable bounds are valid for each $Lk-$level type.

Heterogeneity in bidding functions and beliefs: $k-$rationalizability only describes a sharp upper bound for rationalizable bids. A population of $k-$rationalizable bidders can present a very rich heterogeneity in bidding functions; for instance, some of them may be using the BNE bidding functions while others may be bidding at the boundary. Heterogeneity in biding functions arises from heterogeneity in beliefs.

2.3 Extensions to asymmetric bidders

The setting considered in B-S assumes a collection of n symmetric bidders, but their approach can be extended to allow for asymmetries. Of particular interest would be an environment with a finite number R of *observable* types. Suppose bidders of type r have a distribution of private values $F_r : [0, \overline{v}] \to [0, 1]$ (with common support across all types) and that IPV holds. Maintain the assumption that each F_r is common knowledge. A bidder may have different beliefs (conjectures) about other types. Beliefs for bidder i is a collection of R measures $\{\mu_{ir}\}_{r=1}^{R}$. The corresponding collection of upper bounds $\left\{\overline{B}_{ir,k}(\cdot)\right\}_{r=1}^{R}$, where the subscript ir refers to beliefs held by bidder i about bidders of type r. This may have the potential of generalizing the notion of $k-$rationalizability to something I call $(k_1, \ldots, k_R)-$rationalizability, described by an $R \times R$ matrix of cross-type conjectures about k.

While a full extension of the results in B-S to this general setting is beyond the scope of my paper, I can consider the simple case of auctions with two players ($n = 2$) drawn from two possible types[8]: s (for a small firm) and ℓ (for a large firm). The distributions of private values of different types can be different (F_s is not necessarily equal F_ℓ). F_s and F_ℓ are common knowledge[9]. I do not impose any assumptions on stochastic dominance between F_s and F_ℓ.

Consider type s player. Her private value v_s is independently drawn from the distribution F_s. Her beliefs can be type-specific, denoted as μ_{ss} when her opponent is of type ss and $\mu_{s\ell}$ when her opponent is of type ℓ. Suppose μ_{ss} and $\mu_{s\ell}$ are consistent with k_{ss} and $k_{s\ell}$ steps of elimination of bids that are never best response (define $v_\ell, F_\ell, \mu_{\ell\ell}, \mu_{\ell s}, k_{\ell\ell}$ and $k_{\ell s}$ similarly). I still maintain the assumption that bidders assume that nonzero bids can win with positive probability. My setup allows for different types of players to have misaligned higher-order beliefs about their opponents even in terms of the number of steps they perform, i.e. $k_{s\ell} \neq k_{\ell s} - 1$. This inconsistency of beliefs can become an important source of profit loss for auction participants.

If there are two different types of players s and ℓ in the population, there can be three

[8] This setting is relevant to my empirical illustration with USFS Timber auctions.

[9] The assumptions in B-S on beliefs and behavior imply that players do not use weakly dominated bids. In a second-price auction with private values and any number of participants, this implies that each player bids his valuation. Thus, participating in second-price auctions for the same type of good would allow bidders to learn the distribution of valuations across all types, but it would provide no information about the strategies and beliefs in first-price auctions. Two-thirds of observations in the data set of my empirical illustration are ascending auctions

configurations of types of players in an auction with $n = 2$ players: $(s, s), (s, \ell)$ and (ℓ, ℓ). The configurations (s, s) and (ℓ, ℓ) are just symmetric cases considered in the previous section. I focus on the auction with two different types (s, ℓ). As in the symmetric case, the expected payoffs of bidding b, for bidders of type s and ℓ with private values v_s and v_ℓ respectively are

$$\pi(b, v_s; \mu_{s\ell}) = \int_{B_\ell} (v_s - b)\mathbf{P}[b_\ell \leq b]\mu_{s\ell}(db_\ell) \quad \text{for bidder } s.$$

Also assume that F_s and F_ℓ have the same support and (2.1) is satisfied for both cdfs. Let $\overline{B}_{\ell s,k}(\cdot)$ denote type s's $k-$rationalizable belief about the upper bound for bidding functions by the opponent ℓ. Then,

$$\overline{B}_{\ell s,1}(v_\ell) = v_\ell, \quad \overline{B}_{s\ell,k+1}(v_s) = v_s - \pi^*(v_s; \overline{B}_{\ell s,k}(\cdot)), \quad k = 1, 2, \dots, \text{ where}$$
$$\pi^*(v_s; \overline{B}_{\ell s,k}(\cdot)) = \sup_{b \geq 0} \left\{(v_s - b)\mathbf{P}\left[\overline{B}_{\ell s,k}(v_\ell) \leq b\right]\right\} = \sup_{b \geq 0} \left\{(v_s - b)\int_v 1\left[\overline{B}_{\ell s,k}(v) \leq b\right] dF_\ell(v)\right\}.$$

The upper bound for bidder of type ℓ can be defined in the same replacing s with ℓ and vice versa. The assumption that both F_s and F_ℓ have the same support would prevent the possibility of the $k-$rationalizable bounds may have flat parts. All other properties would follow.

3 An Econometric test for k-rationalizability

In this section I consider symmetric case for simplicity. Since there is only one type of bidders, their beliefs are consistent with the same.

3.1 k-rationalizability as a stochastic dominance restriction

The main goal of my paper is to test whether bidding behavior in a population is consistent with the implications of $k-$rationalizable behavior. Take any bidder i and let $b_i(\cdot)$ denote her bidding function. Denote $b_i(v_i)$ simply as b_i (the actual bid submitted by i). Fix k. Under the assumptions (2.1) and (2.2), bidding behavior is consistent with $k-$rationalizability only if $b_i \leq \overline{B}_k(v_i)$ w.p.1. for each bidder i[10]. This implies a *first-order stochastic dominance* condition:

$$\mathbf{P}\left(\overline{B}_k(v_i) \leq t\right) \leq \mathbf{P}\left(b_i \leq t\right) \quad \forall\, t.$$

I focus on the setting described in Section 2. A crucial property of the $k-$rationalizable bounds is that they are continuous, *strictly increasing* and therefore invertible. So, for all $t \in [\underline{b}, \overline{b}]$, the equation $\overline{B}_k(v) = t$ has a unique solution in v. Denote this solution as $\overline{v}_k(t)$. Then, the above

[10]If all bidders have the same beliefs, the statement is "if and only if".

inequality can be expressed as

$$\mathbf{P}\left(v_i \leq \overline{v}_k(t)\right) \leq \mathbf{P}\left(b_i \leq t\right) \quad \forall\, t \in [\underline{b}, \overline{b}]^{11} \tag{3.1}$$

I could base a test for $k-$rationalizable bidding in the population on the stochastic dominance condition described in (3.1) if I knew (or could identify):

1. The distribution of values.

2. The distribution of bids

Next, I describe a testing procedure for cases where the distribution of values, F_0, is nonparametrically identified from auxiliary data. Specifically, I assume the existence of bidding data on ascending auctions of the same object and the same population of bidders that participate in the first-price auctions.

3.2 A nonparametric test for k-rationalizability with auxiliary data from ascending auctions

The stochastic dominance restriction in (3.1) could be the basis of a nonparametric test for $k-$rationalizability if the distribution of bidders' values were nonparametrically identified. Under the conditions I describe below, F_0 is identified if I have access to transaction prices from ascending (English) auctions for the same type of object and the same population of bidders. I assume the following about the auctions observed.

Assumption 3.1. The same population of bidders participate in both ascending auctions and first price auctions for the same type of object. In any auction with n participants, bidders' private values are independently and identically distributed conditional on a vector of observable auction characteristics X. I assume that the number of participants, n is included in X. Denote this parent distribution as $F_0(\cdot|X)$ and the support[12] $[\underline{v}, \overline{v}]$.

Remark 3.1.

(i) While I assume that $F_0(\cdot|X)$ is the same in both auction formats, I do not require this for the marginal distribution of X, although my results require at least an overlap in the support of the marginal distribution of X in ascending and first-price auctions. Assumption 3.1 basically presupposes that, conditional on the observable characteristics X of the auction, bidders' valuations are not affected by the auction format, a reasonable assumption in light of the private-values environment I focus on.

[11]The support of bids can depend on X and be the type $[\underline{b}(X), \overline{b}(X)]$

[12]In all the results that follow, the support can depend on X and be of the type $[\underline{v}(X), \overline{v}(X)]$.

(ii) A natural model of bidder participation would be the one described in Athey et al. (2011) (see p.236-237). First, all potential bidders observe the auction characteristics X. Then they compare expected profit of participation in the auction with the cost of entry. Potential bidders decide whether to participate simultaneously and may use mixed strategies. Next, participating bidders observe their private values. The equilibrium number of participants n is a function of X and possibly some unobserved heterogeneity ξ that is independent of the private values conditional on X (ξ can be a randomization rule or an equilibrium selection mechanism). This results in a model of exogenous participation where v and n are independent conditional on X.

(iii) Since X can potentially contain information about other bidders' values, I assume that in the first-price auctions (where beliefs are relevant), bidders condition their beliefs on X.

Assumptions (2.1) and (2.2) rule out weakly dominated bids and pin down bidding behavior in second-price auctions. The following result holds (see p.40 in Battigalli and Siniscalchi (2003)):

Result 3.1. *In the second-price auctions bidders submit their private values.*

In my data set, I observe ascending auctions not second-price auction. Thus, I have to make an assumption on the outcome in ascending auctions.

Assumption 3.2. The transaction price observed in the ascending auctions corresponds to the maximum between the reserve price and the second-highest bidder's valuation.

Assumption 3.2 corresponds to the dominant-strategy equilibrium of a button auction version of an ascending auction, but it would also hold (within one bid increment) in the incomplete model of ascending auctions considered in Haile and Tamer (2003) as long as "jump bids" are not observed at the end of an auction[13].

Assumption 3.3.

1. The econometrician observes a sample of L_2 ascending auctions: iid draws $(P_i, X_i)_{i=1}^{L_2}$, where P_i denotes the transaction price (winning bid) in the i^{th} auction and X_i is the vector of observable auction characteristics described in Assumption 3.1. Let $F_{2,P|X}(\cdot|X)$ denote the conditional cdf of P given X in the ascending auctions and let $S_{2,X}$ denote the support of X.

2. The econometrician observes a sample of L_1 first-price sealed-bid auctions of the same type of good with the same population of bidders. $(X_j)_{j=1}^{L_1}$ denotes the corresponding iid sample of observable auction characteristics (X_j includes n_j, the number of participants). Denote $S_{1,X}$ the support of X in first-price auctions. For the j^{th} auction I observe $\left(b_j^i\right)_{i=1}^{n_j}$, the collection of all bids submitted[14] by the n_j bidders. Conditional on $X_j = x$ (with $n_j = n \geq 1$), bids $\left(b_j^i\right)_{i=1}^{n_j}$

[13] Jump bidding occurs when observed changes in bids are above the minimal allowed increment. The presence of jump bids is studied empirically by comparing the relative difference between the two highest bids. I did not find evidence of jump bidding in my data

[14] Under the assumption of symmetric bidders, it is enough to observe transaction price (or a randomly drawn bid) for each auction.

are assumed to be identically distributed[15], with cdf $G_1(\cdot|X)$. Specifically,

$$E\left[\frac{1}{n_j}\sum_{i=1}^{n_j} 1\left[b_j^i \leq t\right] \middle| X_j = x\right] = E\left[1\left[b_j^i \leq t\right]\middle| X_j = x\right] \equiv G_1(t|x) \quad \forall\, j = 1, \ldots, L_1.$$

Remark 3.2. 1. Under the assumption of symmetric bidders, it is enough to observe transaction price (or a randomly drawn bid) for each sealed-bid auction.

2. The most natural way to interpret the restriction that bids $\left(b_j^i\right)_{i=1}^{n_j}$ are identically distributed conditional on X_j is that bids can be written as $b_j^i = b(v_j^i; X_j, \mu(v_j^i; X_j, \xi_j^i))$[16], where the function $b(\cdot)$ is the same across bidders[17], v_j^i denotes i's valuation, ξ_j^i is unobserved heterogeneity that determines i's beliefs and $(v_j^i, \xi_j^i)_{i=1}^{n_j}$ are identically distributed conditional on X_j. Note that this allows heterogeneous beliefs as well as correlation between values and beliefs.

Notation: From now on, I use the subscripts '1' and '2' to denote functionals derived from the population of *first-price auctions* and *ascending auctions* respectively. For instance, $F_{1,X}(\cdot)$ and $F_{2,X}(\cdot)$ denote the cdf of X in first-price and asceding auctions, respectively. Then $E_{1,X}\left[g(X)\right] = \int g(x)dF_{1,X}(x)$, $E_{2,X}\left[g(X)\right] = \int g(x)dF_{2,X}(x)$, etc.

Consider an ascending auction with n bidders. Let $V_{1:n} \leq V_{2:n} \leq \cdots \leq V_{n-1:n} \leq V_{n:n}$ denote the order statistics of their values. Assumption 3.2 implies that, in the absence of a reserve price, the transaction price in this auction corresponds to $V_{n-1:n}$. For $s, t \in [0,1]$ and $n \geq 2$, let

$$\Omega_n(s;t) = t - ns^{n-1} + (n-1)s^n.$$

Take $n \geq 2$ and $x \in S_{2,X}$ (with the component in x corresponding to number of bidders fixed at n). Using the properties of order statistics of iid random variables (David and Nagaraja (2003)), Assumption 3.1 implies that, for any v, the cdf $F_0(v|x)$ is given by the solution, in s, to the equation

$$\Omega_n\left(s; F_{2,P|X}(v|x)\right) = 0$$

I focus the following analysis on the case with no binding reserve price, so $F_0(v|x)$ is nonparametrically identified from $F_{P|X}(v|x)$ through the relationship

$$\Omega_n\left(F_0(v|x); F_{2,P|X}(v|x)\right) = 0, \quad \forall\, v \in [\underline{v}, \overline{v}], x \in S_{2,X} \tag{3.2}$$

[15] I do not allow for unobserved heterogeneity at the auction level. Bids $\left(b_j^i\right)_{i=1}^{n_j}$ can be correlated conditional on X_j, because they are function not only of iid private values, but also of beliefs that can be correlated conditional on X, perhaps through a publicly observed signal that affects beliefs but not values.

[16] Apart from v, the only source of unobserved heterogeneity is beliefs μ. ξ can be of any dimension, even a function.

[17] In the asymmetric case with more than one type of players, bidding function b and belief μ can be indexed by the type t, i.e. $b_{j,t}^i = b_t(v_j^i; X_j, \mu_t(v_j^i; X_j, \xi_j^i))$, where (v_j^i, ξ_j^i) is drawn from the distribution that can depend on t.

Also, with bidders' beliefs conditioned on X, the $k-$rationalizable bids, conditional on v_i and $X = x$ become[18]

$$\overline{B}_1(v_i|x) = v_i, \quad \overline{B}_{k+1}(v_i|x) = v_i - \pi^*(v_i; \overline{B}_k(\cdot|x), x), \quad k = 1, 2, \ldots, \text{ where} \tag{3.3}$$

$$\pi^*(v_i; \overline{B}_k(\cdot|x), x) = \underset{b\geq 0}{sup}\left\{(v_i - b)\boldsymbol{P}\left[\overline{B}_k(v_j|x) \leq b \;\forall\; j \neq i \;\middle|X = x\right]\right\} =$$
$$= \underset{b\geq 0}{sup}\left\{(v_i - b)\left(\int_{\underline{v}}^{\overline{v}} 1\left[\overline{B}_k(v|x) \leq b\right] f_0(v|x)dv\right)^{n-1}\right\} = \underset{b\geq 0}{sup}\left\{(v_i - b)\left(F_0\left(\overline{v}_k(b|x)|x\right)\right)^{n-1}\right\},$$

where, for any $t \in [\underline{b}, \overline{b}]$, $\overline{v}_k(t|x)$ is the (unique) solution, in v, to $\overline{B}_k(v|x) = t$.

The stochastic dominance condition in (3.1) becomes

$$F_0(\overline{v}_k(t|x)|x) \leq G_1(t|x) \quad \forall\; t \in [\underline{b}, \overline{b}], x \in S_{2,X} \cap S_{1,X} \tag{3.4}$$

3.2.1 An econometric test

In this section, I describe an econometric procedure to test $k-$rationalizability based on the stochastic dominance condition (3.4). To construct a test that is not conservative, it is useful to have a procedure that takes into account the properties of the *contact sets* (the regions of (t, x) where the inequalities (3.4) are binding). This helps to avoid conservative tests that use critical values based on so-called "least-favorable configurations" (typically corresponding to the case where the inequalities are binding w.p.1). The procedures in Lee et al. (2014), and Aradillas-Lopez et al. (2016) explicitly take into account the properties of contact sets. Furthermore, both approaches are computationally attractive because they are based on easy-to-compute L_p-statistics, and they have asymptotically pivotal features. However, while Lee et al. (2014) require a direct estimator of the contact sets. The approach in Aradillas-Lopez et al. (2016) relies on a tuning parameter which produces test-statistics that asymptotically adapts to the properties of the contact sets without the need to estimate them. I describe the details of my testing procedure next.

Testing range

Estimating the contact sets directly can be a computationally challenging task especially in the presence of conditioning variables with rich (and unknown) support. Furthermore, since the contact sets themselves are not the object of interest, I will use the type of approach in Aradillas-

[18] With a nonbinding reserve price (i.e, $r \leq \underline{v}$), w.p.1
$$\underset{b\geq 0}{sup}\left\{(v_i - b)\mathbf{P}\left[\overline{B}_k(v_j|x) \leq b \;\forall\; j \neq i \;\middle|X = x\right]\right\} = \underset{b\geq r}{sup}\left\{(v_i - b)\mathbf{P}\left[\overline{B}_k(v_j|x) \leq b \;\forall\; j \neq i \;\middle|X = x\right]\right\}.$$

López et al. (2016) for my testing procedure, by using a tuning parameter that produces a test-statistic that adapts to the properties of the contact sets asymptotically without the need to estimate them. My test relies on nonparametric estimators of the functionals in (3.4). Since I need these estimators to have certain uniform asymptotic properties, I first choose a *testing range* for the values of (x,t) over which I test (3.4).

Assumption 3.4. Denote the support of bids as $[\underline{b},\bar{b}]$. Let $\text{int}(A)$ denote the interior of the set A. Let $S_X = S_{1,X} \cap S_{2,X}$ and assume that $\text{int}(S_{1,X}) \cap \text{int}(S_{2,X})$ is nonempty. Let $\mathscr{X}$ be a compact subset of $\text{int}(S_{1,X}) \cap \text{int}(S_{2,X})$, and let $\mathscr{B}$ be a compact subset of $int([\underline{b},\bar{b}])$. Define

$$\mathscr{V}_k = \{v\colon\ v = \overline{v}_k(b|x) \text{ for some } b \in \mathscr{B} \text{ and } x \in \mathscr{X}.\}.$$

These sets are assumed to satisfy the following conditions:

(i) $0 < \underline{c} \leq G_1(b|x) \leq \bar{c} < 1$ for all $b \in \mathscr{B}$, $x \in \mathscr{X}$. Also, $f_{2,X}(x) \geq \underline{f} > 0$ and $f_{1,X}(x) \geq \underline{f} > 0$ for all $x \in \mathscr{X}$, where $f_{2,X}(\cdot)$ and $f_{1,X}(\cdot)$ denote the densities of X in the population of ascending and first-price auctions respectively.

(ii) $0 < \underline{c} \leq F_0(v|x) \leq \bar{c} < 1$ and $0 < \underline{c} \leq F_{2,P|X}(v|x) \leq \bar{c} < 1$ for all $v \in \mathscr{V}_k$, $x \in \mathscr{X}$.

I rewrite restriction (3.4) in terms of a mean-zero condition. For a given $t \in [\underline{b},\bar{b}]$, $x \in S_{2,X}$ and $k \geq 1$, define

$$\phi_k(t|x) \equiv F_0(\overline{v}_k(t|x)|x) - G_1(t|x).$$

Recall that $G_1(\cdot|X)$ denotes the conditional cdf of bids given X in first-price auctions. If bids are $k-$rationalizable, one must have $\phi_k(t|x) \leq 0$ for all $x \in S_{1,X} \cap S_{2,X}$ and all $t \in [\underline{b},\bar{b}]$; in particular $\phi_k(t|x) \leq 0$ for all (x,t) such that $x \in \mathscr{X}$ and $t \in \mathscr{B}(x)$. **This is the basis of my test.**

Denote $(V)_+ \equiv \max\{V, 0\}$. For each $x \in \mathscr{X}$, let $\mathscr{B}(x) \subseteq \mathscr{B}$, and let Q be a pre-specified probability measure for t conditional on x satisfying $\int_{t\in\mathscr{B}(x)} dQ(t|x) = 1$ for all $x \in \mathscr{X}$. Let $W_{\mathscr{X}}$ be a nonnegative weighting function that satisfies $W_{\mathscr{X}}(x) > 0$ if $x \in \mathscr{X}$ and $W_{\mathscr{X}}(x) = 0$ otherwise. Let

$$\begin{aligned} \Lambda_k(x) &\equiv \int_{t\in\mathscr{B}(x)} (\phi_k(t|x))_+ \, dQ(t|x), \\ \mathcal{T}_k &\equiv E_{2,X}\left[\Lambda_k(X)\cdot W_{\mathscr{X}}(X)\right] + E_{1,X}\left[\Lambda_k(X)\cdot W_{\mathscr{X}}(X)\right] \end{aligned} \tag{3.5}$$

Note that $\mathcal{T}_k \geq 0$. Bidders are using $k-$rationalizable bids in first-price auctions only if $\mathcal{T}_k = 0$. On the other hand, $\mathcal{T}_k > 0$ implies necessarily a violation of $k-$rationalizability. **My test will be a one sided test for the null hypothesis $\boldsymbol{H_0 : \mathcal{T}_k = 0}$ against the alternative $\boldsymbol{H_1 : \mathcal{T}_k > 0}$.**

A nonparametric estimator for $\mathcal{T}_k$

I propose a nonparametric test-statistic along the lines of Lee et al. (2013), Lee et al. (2014) and Aradillas-Lopez et al. (2016). For this reason I need to add restrictions that ensure that the proposed nonparametric estimators have the desired asymptotic properties. The following assumption describes the main features about the data.

Assumption 3.5. The vector of auction characteristics X is partitioned as $X \equiv \left(X^c, X^d\right)$, where X^c are those components assumed to be continuously distributed and X^d are discrete (the number of participants in the auction is included in X^d). Let $c \equiv dim(X^c)$ denote the number of continuously distributed elements in X. Let $f_{2,X^c}(\cdot)$ and $f_{1,X^c}(\cdot)$ denote the density of X^c in the population of ascending and first-price auctions, respectively with $p_{2,X^d|X^c}(x^d|x^c) = \boldsymbol{P}_{2,X}(X^d = x^d|X^c = x^c)$ and $p_{1,X^d|X^c}(x^d|x^c) = \boldsymbol{P}_{1,X}(X^d = x^d|X^c = x^c)$. Then $f_{2,X}(x) = p_{2,X^d|X^c}(x^d|x^c)f_{2,X^c}(x^c)$ and $f_{1,X}(x) = p_{1,X^d|X^c}(x^d|x^c)f_{1,X^c}(x^c)$ describes the joint density of X in each case.

I treat bids b (in first-price auctions), transaction price P (in ascending auctions) and X^c (the continuous components in X) as continuously distributed. I use kernel-based nonparametric estimators with a multiplicative kernel of the type $K(\psi) = \prod_{i=1}^{c} k(\psi_i)$ for X^c, where the individual kernel $k : \mathbb{R} \to \mathbb{R}$ is a symmetric function around zero with additional properties that are described below. I use separate sets of bandwidths for x^c, b and P. These are denoted as h_{x,L_1} (for X^c) and h_{b,L_1} (for b) in the sample of first-price auctions, and as h_{x,L_2} (for X^c) and h_{p,L_2} (for P) in the sample of ascending auctions. The exact properties of the kernel function k and the bandwidths are explained in Assumption 3.9.

3.2.2 Nonparametric estimators from the sample of ascending auctions

The distribution of values F_0 and the $k-$rationalizable bounds are nonparametrically estimated from the ascending auctions sample. For a given $x \equiv (x^c, x^d)$ and $p \in \mathbb{R}$ let

$$\mathcal{H}(X_i - x; h_{x,L_2}) = K\left(\frac{X_i^c - x^c}{h_{x,L_2}}\right) 1\left[X_i^d = x^d\right], \qquad \widehat{f}_{2,X}(x) = \left(L_2 h_{x,L_2}^c\right)^{-1} \sum_{i=1}^{L_2} \mathcal{H}(X_i - x; h_{x,L_2}),$$

$$\widehat{f}_{2,(P,X)}(p,x) = \left(L_2 h_{x,L_2}^c\right)^{-1} \sum_{i=1}^{L_2} \frac{1}{h_{p,L_2}} k\left(\frac{P_i - p}{h_{p,L_2}}\right) \mathcal{H}(X_i - x; h_{x,L_2}), \qquad \widehat{f}_{2,P|X}(p|x) = \frac{\widehat{f}_{2,(P,X)}(p,x)}{\widehat{f}_{2,X}(x)},$$

$$\widehat{F}_{2,P|X}(p|x) = \int_{-\infty}^{p} \widehat{f}_{2,P|X}(t|x)dt = \frac{\left(L_2 h_{x,L_2}^c\right)^{-1} \sum_{i=1}^{L_2} \left(\int_{t=-\infty}^{p} \frac{1}{h_{p,L_2}} k\left(\frac{P_i - t}{h_{p,L_2}}\right) dt\right) \cdot \mathcal{H}(X_i - x; h_{x,L_2})}{\widehat{f}_{2,X}(x)}.$$

My choice of estimator for $\widehat{F}_{2,P|X}(\cdot|x)$ helps simplify the asymptotic analysis because its argument is itself a nonparametric estimator. Note that using Leibniz rule,

$$\frac{d\widehat{F}_{2,P|X}(p|x)}{dp} = \frac{\left(Lh^c_{x,L_2}\right)^{-1}\sum_{i=1}^{L_2}\frac{1}{h_{p,L_2}}k\left(\frac{P_i-p}{h_{p,L_2}}\right)\cdot\mathcal{H}(X_i-x;h_{x,L_2})}{\widehat{f}_{2,X}(x)} = \widehat{f}_{2,P|X}(p|x).$$

For a given (v,x) and an ascending auction with n bidders (the component in x^d corresponding to number of bidders is therefore fixed at n), I estimate conditional distribution of private values $F_0(v|x)$ using (3.2) after I replace $F_{2,P|X}(v|x)$ with $\widehat{F}_{2,P|X}(v|x)$. My estimator $\widehat{F}_0(v|x)$ is the solution, in s, to the equation $\Omega_n\left(s;\widehat{F}_{2,P|X}(v|x)\right)=0$. Therefore $\widehat{F}_0(v|x)$ is defined implicitly by the equation

$$\Omega_n\left(\widehat{F}_0(v|x);\widehat{F}_{2,P|X}(v|x)\right)=0. \tag{3.6}$$

Estimation of the $k-$rationalizable bounds

I estimate the bounds with sample of ascending auctions analogs to the construction in (3.3). For a given v and x, I have

$$\widehat{\overline{B}}_1(v|x)=v, \quad \widehat{\overline{B}}_{k+1}(v|x) \ = v-\widehat{\pi}^*(v;\overline{B}_k(\cdot|x),x), \quad k=1,2,\ldots, \quad \text{where} \tag{3.7}$$

$$\widehat{\pi}^*(v;\widehat{\overline{B}}_k(\cdot|x),x)=\underset{b\geq 0}{sup}\left\{(v-b)\left(\widehat{F}_0\left(\widehat{\overline{v}}_k(b|x)|x\right)\right)^{n-1}\right\},$$

where $\widehat{\overline{v}}_k(t|x)$ solves (in v) the condition $\widehat{\overline{B}}_k(v|x)=t$ (with $t\in[\underline{b},\overline{b}]$). The least-favorable $k-$rationalizable estimated expected payoff $\widehat{\pi}^*(v;\widehat{\overline{B}}_k(\cdot|x),x)$ is constructed as follows. Let

$$\widehat{b}^*_k(v|x)=\underset{b\geq 0}{argmax}\left\{(v-b)\left(\widehat{F}_0\left(\widehat{\overline{v}}_k(b|x)|x\right)\right)^{n-1}\right\}, \quad \widehat{\overline{v}}^*_k(v|x) \ \equiv \widehat{\overline{v}}_k(\widehat{b}^*_k(v|x)|x),$$

Then, $\widehat{\pi}^*(v;\widehat{\overline{B}}_k(\cdot|x),x)=(v-\widehat{b}^*_k(v|x))\left(\widehat{F}_0\left(\widehat{\overline{v}}^*_k(v|x)|x\right)\right)^{n-1}$.

3.2.3 Nonparametric estimators from the sample of first-price auctions

The distribution of bids G_1 in first-price auctions is estimated nonparametrically from this sample. Fix x and let n denote the element in x corresponding to the number of bidders. I estimate $G_1(\cdot|x)$ as follows.

$$\widehat{f}_{1,X}(x)=\left(L_1h^c_{x,L_1}\right)^{-1}\sum_{j=1}^{L_1}\mathcal{H}(X_j-x;h_{x,L_1}),$$

$$\widehat{g}_{1,(b,X)}(b,x) = \left(L_1 h_{x,L_1}^c\right)^{-1} \sum_{j=1}^{L_1} \mathcal{H}(X_j - x; h_{x,L_1}) \cdot (n \cdot h_{b,L_1})^{-1} \sum_{i=1}^{n} k\left(\frac{b_j^i - b}{h_{b,L_1}}\right),$$

$$\widehat{g}_{1,b|X}(b|x) = \frac{\widehat{g}_{1,(b,X)}(b,x)}{\widehat{f}_{1,X}(x)},$$

$$\widehat{G}_1(b|x) = \int_{-\infty}^{b} \widehat{g}_{1,b|X}(t|x)dt = \frac{\left(L_1 h_{x,L_1}^c\right)^{-1} \sum_{j=1}^{L_1} \mathcal{H}(X_j - x; h_{x,L_1}) \cdot (n \cdot h_{b,L_1})^{-1} \sum_{i=1}^{n} \left(\int_{-\infty}^{b} k\left(\frac{b_j^i - t}{h_{b,L_1}}\right) dt\right)}{\widehat{f}_{1,X}(x)}.$$

3.2.4 Estimation of $\mathcal{T}_k$

To construct the estimator of $\mathcal{T}_k$, I combine the nonparametric estimators obtained from both samples, as follows. First, for a given k and (t,x), let

$$\widehat{\phi}_k(t|x) = \widehat{F}_0(\widehat{\overline{v}}_k(t|x)|x) - \widehat{G}_1(t|x).$$

Let $\gamma_{L_1} \to 0$ and $\gamma_{L_2} \to 0$ be two tuning parameters (positive bandwidth sequences) indexed by the sample sizes L_1 and L_2, respectively. Let

$$\widehat{\Lambda}_{k,q}(x) = \int_{t \in \mathscr{B}(X)} \widehat{\phi}_k(t|x) \cdot 1\left[\widehat{\phi}_k(t|x) \geq -\gamma_{L_q}\right] dQ(t|x), \quad q = 1,2.$$

The rate-of-convergence restrictions for γ_{L_1} and γ_{L_2} are described in Assumption 3.9, below. For a testing range $\mathscr{X}$, $\mathscr{B}$ and for the weighting functions Q and $W_{\mathscr{X}}$, I estimate

$$\widehat{\mathcal{T}}_k = \frac{1}{L_2} \sum_{i=1}^{L_2} \widehat{\Lambda}_{k,2}(X_i) W_{\mathscr{X}}(X_i) + \frac{1}{L_1} \sum_{j=1}^{L_1} \widehat{\Lambda}_{k,1}(X_j) W_{\mathscr{X}}(X_j), \tag{3.8}$$

The use of the tuning parameters γ_{L_1} and γ_{L_2} generates asymptotic properties for $\widehat{\mathcal{T}}_k$ that automatically adapt to the properties of the so-called *contact sets*, which for a given k are defined as

$$\{(t,x)\colon\ x \in \mathscr{X}, t \in \mathscr{B}(x),\ \phi_k(t|x) = 0\}.$$

The contact sets would correspond to regions of X where bidders bid at the boundary of the $k-$rationalizable bids.

Assumption 3.6. For some integer $M \geq 1$ (restrictions on it are described below), the following holds: evaluated at any $v \in (\underline{v}, \overline{v})$, $b \in (\underline{b}, \overline{b})$ and $x \in \mathscr{X}$, $f_0(v|x)$, $f_{2,P|X}(v|x)$, $g_{1,b|X}(b|x)$, $f_{2,X}(x)$ and $f_{1,X}(x)$ are $M-$times differentiable with respect to x^c (the continuous elements of x) with bounded derivatives.

An additional assumption is needed regarding the stochastic properties of $\phi_k(t|X)$. Obviously, I must allow the possibility for $\phi_k(t|X)$ to have a point-mass at zero at least over some range of values $t \in \mathscr{B}(X)$ (which occurs if some bidders in the population are bidding exactly at the $k-$rationalizable upper bound), but I need to impose a reasonably mild restriction on the density of $\phi_k(t|X)$ to the left of zero (in an interval of the type $[-b, 0)$). Basically the only restriction I impose is that the density is finite over such interval (conditional on $X \in \mathscr{X}$). I describe the condition next.

Assumption 3.7. There exists $\underline{\gamma} > 0$ and $D < \infty$ such that, for all $t \in \mathscr{B}$,

$$\mathbf{P}_{q,X}\left(-c \leq \phi_k(t|X) < 0 \middle| X \in \mathscr{X}\right) \leq D \cdot c \quad \forall\, 0 < c \leq \underline{\gamma}, \quad q = 1, 2.$$

The asymptotic properties of my proposed test-statistic are partially influenced by those of two empirical processes $\nu_{L_2}(\cdot)$ and $\nu_{L_1}(\cdot)$ indexed over $\mathscr{B}$, where

$$\nu_{L_q}(t) = \frac{1}{\sqrt{L_q}} \sum_{i=1}^{L_q} \Big(1\left[-2\gamma_{L_q} \leq \phi_k(t|X_i) < 0\right] W_{\mathscr{X}}(X_i) - E_{q,X}\left[1\left[-2\gamma_{L_q} \leq \phi_k(t|X_i) < 0\right] W_{\mathscr{X}}(X_i)\right] \Big), q = 1, 2.$$

The following condition suffices to ensure that these processes are manageable in the sense of Pollard (1990). Let $\mathcal{C}$ be a class of subsets of a set $\mathcal{X}$. For any set X_0 of n points in $\mathcal{X}$ let $\mathcal{C}(X_0) = \{X_0 \cap C \colon C \in \mathcal{C}\}$. Notice that $\#\mathcal{C}(X_0)$ corresponds to the number of subsets of X_0 that can be picked out by some C in $\mathcal{C}$. It is said that the class $\mathcal{C}$ shatters X_0 if $\#\mathcal{C}(X_0) = 2^n$. The Vapnik-Cervonenkis dimension (or VC dimension) $(V(\mathcal{C}))$ of $\mathcal{C}$ is defined as the largest n such that there exists *some* X_0 with n points that is shattered by $\mathcal{C}$. Class $\mathcal{C}$ is a VC-class of sets if $V(\mathcal{C}) < \infty$ (see Section 9.1.1 in Kosorok (2008)).

Assumption 3.8. There exists a $\bar{b} > 0$ such that

$$\left\{\{x \in \mathscr{X} \colon -\gamma \leq \phi_k(t|x) < 0\} \colon t \in \mathscr{B},\, 0 < \gamma \leq \bar{b}\right\}$$

is a VC class of sets.

Next I describe the restrictions governing rates of convergence of the tuning parameters, as well as the kernel function and a condition related to how L_1 and L_2 grow asymptotically.

Assumption 3.9.

1. The multivariate kernel K is multiplicative, with $K(u) = \prod_{j=1}^{c} k(u_j)$ for $u = (u_1, \ldots, u_c)$. The marginal kernel k is a bounded function, symmetric around zero and it has support over a compact interval $[-S, S]$. It is a bias-reducing kernel of order M (where M is introduced in Assumption 3.6), meaning that it satisfies

$$\int_{-S}^{S} k(u)du = 1, \quad \int_{-S}^{S} u^j k(u)du = 0,\, j = 1, \ldots, M-1, \quad \int_{-S}^{S} |u|^M k(u)du < \infty.$$

$k(\cdot)$ is differentiable almost everywhere on $[-S, S]$ with bounded first derivative $k'(\cdot)$.

2. The sample sizes L_1 and L_2 satisfy a proportionality condition in the limit,

$$\lim_{\substack{L_1 \to \infty \\ L_2 \to \infty}} \left(\frac{L_2}{L_1 + L_2} \right) = d_2 > 0, \quad \lim_{\substack{L_1 \to \infty \\ L_2 \to \infty}} \left(\frac{L_1}{L_1 + L_2} \right) = d_1 > 0$$

3. The bandwidths h_{b,L_1} (for b in first-price auctions) and h_{p,L_2} (for transaction price P in ascending auctions) satisfy:

   (i) $L_1 h_{b,L_1} \longrightarrow \infty$ and $L_1^{1/2} h_{b,L_1} \longrightarrow 0$.

   (ii) $L_2 h_{p,L_2} \longrightarrow \infty$ and $L_2^{1/2} h_{p,L_2} \longrightarrow 0$.

4. For both $q = 1, 2$, the following conditions are satisfied. The convergence rates of the bandwidths $h_{x,L_q} \longrightarrow 0$ and $\gamma_{L_q} \longrightarrow 0$ satisfy the following:

   (i) $L_q^{1/2} h_{x,L_q}^c \gamma_{L_q} \longrightarrow \infty$.

   (ii) $L_q^{1/2+\delta} \gamma_{L_q}^2 \longrightarrow 0$ for some $\delta > 0$.

   (iii) For the M and $\delta > 0$ described above, $L_q^{1/2+\delta} h_{x,L_q}^M \longrightarrow 0$

Note that I require that the auxiliary bandwidth γ_{L_q} converge to zero slower than the rate of convergence of the nonparametric estimators $(L_q h_{x,L_q})^{-1/2}$, but fast enough that the square of γ_{L_q} go to zero faster than $L_q^{-1/2}$. Taken together, my assumptions require that γ_{L_q} converge to zero faster than h_{L_q} (the conditions described imply that $(\gamma_{L_q}/h_{x,L_q}^c) \longrightarrow 0$). I also assume essentially that the sample sizes L_1 and L_2 grow at a proportional rate, meaning that I rule out $L_2/L_1 \longrightarrow 0$ and $L_1/L_2 \longrightarrow 0$. I also impose smoothness restrictions (with respect to the continuous elements in X) that are commonly present in the nonparametric literature, where the degree of smoothness required increases with c (the dimension of X^c). The smallest value of M that can be consistent with my assumptions is $M = 2c + 1$. Finally, note that both h_{b,L_1} and h_{p,L_2} converge to zero faster than h_{x,L_1} and h_{x,L_2}. The only purpose of the first two bandwidths is to "smooth out" the indicator function in the construction of the cdf estimators $\widehat{G}_1$ and $\widehat{F}_{2,P|X}$. The main result for $\widehat{\mathcal{T}}_k$ is the following.

Theorem 3.1. *Let $L \equiv L_1 + L_2$. If Assumptions 3.1-3.9 are satisfied, then for some $\Delta > \frac{1}{2}$,*

$$\widehat{\mathcal{T}}_k = \mathcal{T}_k + \frac{1}{L_1} \sum_{j=1}^{L_1} \psi_{1,k}(\boldsymbol{b}_j, X_j; h_{b,L_1}, h_{x,L_1}) + \frac{1}{L_2} \sum_{i=1}^{L_2} \psi_{2,k}(P_i, X_i; h_{p,L_2}, h_{x,L_2}) + o_p\left(L^{-\Delta}\right).$$

where $\psi_{1,k}$ and $\psi_{2,k}$ are two influence functions satisfying the following conditions:

(i) $E_{1,(X,b)}\left[\psi_{1,k}(\boldsymbol{b},X;h_{b,L_1},h_{x,L_1})\right]=E_{2,(X,P)}\left[\psi_{2,k}(P_i,X_i;h_{p,L_2},h_{x,L_2})\right]=0$.

(ii) *If* $\phi_k(t|x)<0$ *for almost every* $x\in\mathscr{X}$, $t\in\mathscr{B}(x)$ *(i.e, if bids are strictly below the* $k-$*rationalizable bounds almost everywhere over the testing range), then*

$$\psi_{1,k}(\boldsymbol{b},X;h_{b,L_1},h_{x,L_1})=\psi_{2,k}(P_i,X_i;h_{p,L_2},h_{x,L_2})=0 \quad w.p.1$$

(iii) *Let* $E_{1,(X,b)}\left[\left(\psi_{1,k}(\boldsymbol{b},X;h_{b,L_1},h_{x,L_1})\right)^2\right]\equiv\sigma^2_{1_k,L_1}$ *and* $E_{2,(X,P)}\left[\left(\psi_{2,k}(P_i,X_i;h_{p,L_2},h_{x,L_2})\right)^2\right]\equiv\sigma^2_{2_k,L_2}$. *Then, if* $\phi_k(t|x)\geq 0$ *with positive probability over our testing range, then*

$$\lim_{L\to\infty}\sigma^2_{1_k,L_1}\equiv\sigma^2_{1_k}>0 \quad \textit{and} \quad \lim_{L\to\infty}\sigma^2_{2_k,L_2}\equiv\sigma^2_{2_k}>0.$$

Proof: A step-by-step proof is included in Econometric Appendix A (the exact expressions for $\psi_{1,k}(\boldsymbol{b}_j,X_j;h_{b,L_1},h_{x,L_1})$ and $\psi_{2,k}(P_i,X_i;h_{p,L_2},h_{x,L_2})$ can be found in Equation (A.41)). A sketch of the proof is the following. The first part of the proof is to show that, under the assumptions of the theorem, $\widehat{\phi}_k(t|x)$ has a linear representation of the form

$$\begin{aligned}\widehat{\phi}_k(t|x)-\phi_k(t|x)=&\frac{1}{L_2h^c_{x,L_2}}\sum_{i=1}^{L_2}\varphi^{F_{0,k}}\left(P_i,X_i,t,x;h_{p,L_2},h_{x,L_2}\right)-\\&\frac{1}{L_1h^c_{x,L_1}}\sum_{j=1}^{L_1}\varphi^{G_1}\left(\boldsymbol{b}_j,X_j,t,x;h_{b,L_1},h_{x,L_1}\right)+\xi^{F_{0,k}}_{L_2}(t,x)-\xi^{G_1}_{L_1}(t,x),\end{aligned}$$

$$\begin{cases}E_{2,(P,X)}\left[\varphi^{F_{0,k}}\left(P_i,X_i,t,x;h_{p,L_2},h_{x,L_2}\right)\right]=0,\\ E_{1,(\boldsymbol{b},X)}\left[\varphi^{G_1}\left(\boldsymbol{b}_j,X_j,t,x;h_{b,L_1},h_{x,L_1}\right)\right]=0\end{cases} \quad \text{for all } t\in\mathscr{B},\ x\in\mathscr{X}.$$

$$\begin{cases}\underset{\substack{t\in\mathscr{B}\\x\in\mathscr{X}}}{sup}\left|\xi^{F_{0,k}}_{L_2}(t,x)\right|=O_p\left(L_2^{-1/2-\epsilon}\right),\\ \underset{\substack{t\in\mathscr{B}\\x\in\mathscr{X}}}{sup}\left|\xi^{G_1}_{L_1}(t,x)\right|=O_p\left(L_1^{-1/2-\epsilon}\right).\end{cases} \quad \text{for some } \epsilon>0.$$

The expression for $\varphi^{F_{0,k}}$ is derived inductively, starting with $k=2$. From here, the next step is to show that

$$\begin{aligned}\frac{1}{L_1}\sum_{i=1}^{L_1}\widehat{\Lambda}_{k,1}(X_i)W_{\mathscr{X}}(X_i)=&E_{1,X}\left[\Lambda_k(X_i)W_{\mathscr{X}}(X_i)\right]+\frac{1}{L_1}\sum_{i=1}^{L_1}\Big(\Lambda_k(X_i)W_{\mathscr{X}}(X_i)-E_{1,X}\left[\Lambda_k(X_i)W_{\mathscr{X}}(X_i)\right]\Big)\\&+\frac{1}{L_1L_2h^c_{x,L_2}}\sum_{i=1}^{L_1}\sum_{\ell=1}^{L_2}\int_{t\in\mathscr{B}(X_i)}\varphi^{F_{0,k}}\left(P_\ell,X_\ell,t,X_i;h_{p,L_2},h_{x,L_2}\right)1\left[\phi_k(t|X_i)\geq 0\right]dQ(t|X_i)\cdot W_{\mathscr{X}}(X_i)\end{aligned}$$

$$
-\frac{1}{L_1^2 h_{x,L_1}^c}\sum_{i=1}^{L_1}\sum_{m=1}^{L_1}\int_{t\in\mathscr{B}(X_i)}\varphi^{G_1}\left(\boldsymbol{b}_m,X_m,t,X_i;h_{b,L_1},h_{x,L_1}\right)1\left[\phi_k(t|X_i)\geq 0\right]dQ(t|X_i)W_{\mathscr{X}}(X_i)+\widehat{\varpi}_{k,1},
$$

$$
\frac{1}{L_2}\sum_{i=1}^{L_2}\widehat{\Lambda}_{k,2}(X_i)W_{\mathscr{X}}(X_i)=E_{2,X}\left[\Lambda_k(X_i)W_{\mathscr{X}}(X_i)\right]+\frac{1}{L_2}\sum_{i=1}^{L_2}\Big(\Lambda_k(X_i)W_{\mathscr{X}}(X_i)-E_{2,X}\left[\Lambda_k(X_i)W_{\mathscr{X}}(X_i)\right]\Big)
$$

$$
+\frac{1}{L_2^2 h_{x,L_2}^c}\sum_{i=1}^{L_2}\sum_{\ell=1}^{L_2}\int_{t\in\mathscr{B}(X_i)}\varphi^{F_{0,k}}\left(P_\ell,X_\ell,t,X_i;h_{p,L_2},h_{x,L_2}\right)1\left[\phi_k(t|X_i)\geq 0\right]dQ(t|X_i)\cdot W_{\mathscr{X}}(X_i)
$$

$$
-\frac{1}{L_2 L_1 h_{x,L_1}^c}\sum_{i=1}^{L_2}\sum_{m=1}^{L_1}\int_{t\in\mathscr{B}(X_i)}\varphi^{G_1}\left(\boldsymbol{b}_m,X_m,t,X_i;h_{b,L_1},h_{x,L_1}\right)1\left[\phi_k(t|X_i)\geq 0\right]dQ(t|X_i)W_{\mathscr{X}}(X_i)+\widehat{\varpi}_{k,2},
$$

where $\widehat{\varpi}_{k,1}=O_p\left(L^{-1/2-\epsilon}\right)$ and $\widehat{\varpi}_{k,2}=O_p\left(L^{-1/2-\epsilon}\right)$ for some $\epsilon>0$. Note that if $\phi_k(t|x)<0$ almost surely over the testing range (i.e, if the stochastic dominance inequalities from $k-$rationalizability hold *strictly* w.p.1), then $\frac{1}{L_1}\sum_{i=1}^{L_1}\widehat{\Lambda}_{k,1}(X_i)W_{\mathscr{X}}(X_i)=O_p\left(L^{-1/2-\epsilon}\right)$ and $\frac{1}{L_2}\sum_{i=1}^{L_2}\widehat{\Lambda}_{k,2}(X_i)W_{\mathscr{X}}(X_i)=O_p\left(L^{-1/2-\epsilon}\right)$ for some $\epsilon>0$. The final step of the proof comes from the Hoeffding decompositon (see Lemma 5.1.A in Serfling (1980)) of the (generalized) U-statistics that appear on the right-hand side of the above expressions. ■

A test based on Theorem 3.1

I test the null hypothesis that bids by every bidder in the population are consistent with $k-$rationalizability against the alternative that, with positive probability, there exist bidders who submit bids that violate the $k-$rationalizable bounds. My test is based on the stochastic-dominance implications of $k-$rationalizability described in (3.4). The result in Theorem 3.1 guides the construction of a test of $k-$rationalizability:

1. If bids violate the $k-$rationalizable bounds, then

$$
\sqrt{L}\cdot\widehat{\mathcal{T}}_k=\sqrt{L}\cdot\mathcal{T}_k+O_p(1),
$$

   and therefore $\sqrt{L}\cdot\widehat{\mathcal{T}}_k\longrightarrow+\infty$ w.p.1.

2. If bids satisfy the $k-$rationalizable bounds, then:

   (a) If bidders bid strictly below the $k-$rationalizable bounds everywhere over the testing range $\mathscr{X}\times\mathscr{B}(x)$, then

$$
\sqrt{L}\cdot\widehat{\mathcal{T}}_k=o_p(1).
$$

   (b) If bidders bid at the $k-$rationalizable bounds with positive probability over the testing range and the stochastic dominance inequalities are binding, then

$$\frac{\widehat{\mathcal{T}}_k}{\sigma_{k,L_1,L_2}} \xrightarrow{d} N(0,1),$$

where

$$\sigma^2_{k,L_1,L_2} \equiv \frac{\sigma^2_{1_k,L_1}}{L_1} + \frac{\sigma^2_{2_k,L_2}}{L_2}. \tag{3.9}$$

Since I can have $\sigma^2_{k,L_1,L_2} = 0$ if bids are strictly below the $k-$rationalizable bounds a.e over the testing range, in order to construct a test I need to regularize σ^2_{k,L_1,L_2}. This can be done in several ways, but I do it in a way that does not lead to overrejection and allow to avoid estimation of the contact sets that results in a conservative test. Let $\lambda_L \to 0$ be a positive sequence converging to zero very slowly. Specifically, suppose $\lambda_L \cdot L^\epsilon \longrightarrow \infty$ for any $\epsilon > 0$ (this is true, for example, if $\lambda_L \propto (\log L)^{-1}$). Let

$$\bar{t}_k = \frac{\widehat{\mathcal{T}}_k}{\sqrt{\max\left\{\sigma^2_{k,L_1,L_2}, \frac{\lambda_L}{L}\right\}}}$$

Then,

1. If bids violate the $k-$rationalizable bounds,

$$\bar{t}_k = \underbrace{\frac{1}{\sqrt{\max\left\{\left(\left(\frac{L}{L_1}\right)\sigma^2_{1_k,L_1} + \left(\frac{L}{L_2}\right)\sigma^2_{2_k,L_2}\right), \lambda_L\right\}}}}_{\longrightarrow \frac{1}{\sqrt{\left(\frac{1}{d_1}\right)\sigma^2_{1_k} + \left(\frac{1}{d_2}\right)\sigma^2_{2_k}}}, \text{ where } \sigma^2_{1_k} > 0 \text{ and } \sigma^2_{2_k}} \times \underbrace{\sqrt{L}\cdot\widehat{\mathcal{T}}_k}_{\xrightarrow{p} +\infty \text{ w.p.1.}}$$

   Then, $\bar{t}_k \xrightarrow{p} +\infty$ in this case.

2. If bids satisfy the $k-$rationalizable bounds,

   (a) If bidders bid strictly below the bounds over the testing range,

$$\bar{t}_k = o_p\left(\frac{L^{1/2}\cdot L^{-\Delta}}{\lambda_L^{1/2}}\right) = o_p(1), \text{ since } \Delta > 1/2 \text{ and } \lambda_L L^\epsilon \to \infty \ \forall\ \epsilon > 0.$$

   Then, $\bar{t}_k \xrightarrow{p} 0$ in this case.

   (b) If bidders bid at the $k-$rationalizable bounds with positive probability over the testing range, then $\sqrt{L}\cdot\widehat{\mathcal{T}}_k \xrightarrow{d} N\left(0, \frac{1}{d_1}\sigma^2_{1,k} + \frac{1}{d_2}\sigma^2_{2,k}\right)$, and

$$\bar{t}_k = \frac{1}{\sqrt{\max\left\{\left(\left(\frac{L}{L_1}\right)\sigma^2_{1_k,L_1} + \left(\frac{L}{L_2}\right)\sigma^2_{2_k,L_2}\right), \lambda_L\right\}}} \times \sqrt{L}\cdot\widehat{\mathcal{T}}_k \xrightarrow{d} N(0,1)$$

Let $\widehat{\sigma}^2_{k,L_1,L_2}$ be a consistent estimator of σ^2_{k,L_1,L_2} (i.e, $\left|\widehat{\sigma}^2_{k,L_1,L_2} - \sigma^2_{L_1,L_2}\right| \xrightarrow{p} 0$). This can be obtained, for example, with

$$\widehat{\sigma}^2_{k,L_1,L_2} = \left(\frac{1}{L_1}\right) \cdot \left[\frac{1}{L_1}\sum_{j=1}^{L_1} \widehat{\psi}_{k,1}(\boldsymbol{b}_j, X_j; h_{b,L_1}, h_{x,L_1})^2\right] + \left(\frac{1}{L_2}\right) \cdot \left[\frac{1}{L_2}\sum_{i=1}^{L_2} \widehat{\psi}_{k,2}(P_i, X_i; h_{p,L_2}, h_{x,L_2})^2\right].$$

Given the linear representation asymptotic properties, a bootstrap estimator for $\sigma^2_{k,L}$ would also have consistency properties. My test-statistic is

$$\widehat{t}_k = \frac{\widehat{\mathcal{T}}_k}{\sqrt{\max\left\{\widehat{\sigma}^2_{k,L_1,L_2},\ \frac{\lambda_L}{L}\right\}}}. \tag{3.10}$$

Rejection rule: The null and alternative hypotheses are:

H_0 : Bids satisfy the restrictions of $k-$rationalizability w.p.1.

H_1 : Bids violate the restrictions of $k-$rationalizability with positive probability.

For a target significance level α, let $\Phi(z_{1-\alpha}) = 1 - \alpha$, where $\Phi(\cdot)$ is the Standard Normal distribution. In view of the asymptotic properties of $\widehat{t}_k$, the rejection rule is:

$$\text{Reject } H_0 \text{ if and only if } \widehat{t}_k > z_{1-\alpha}.$$

From the asymptotic properties of $\widehat{t}_k$, this rejection rule has the following features:

1. $\lim_{L\to\infty} Pr\,(\text{Falsely rejecting } H_0) \leq \alpha$.
2. $\lim_{L\to\infty} Pr\,(\text{Falsely rejecting } H_0) = \alpha$ if bids lie at the $k-$rationalizable upper bound with positive probability and the stochastic dominance inequalities are binding over our testing range.
3. $\lim_{L\to\infty} Pr\,(\text{Rejecting } H_0) = 1$ if bids violate the $k-$rationalizable upper bounds and the stochastic dominance inequalities are violated with positive probability over our testing range.

Using Standard Normal critical values yields a testing procedure that is computationally easy to implement. However, the linear representation result in Theorem 3.1 also facilitates the analysis of resampling-based methods; in particular, the fact that my results produce exact analytical expressions for the functions $\psi_{k,1}$ and $\psi_{k,2}$ immediately implies that the Multiplier Bootstrap can be used (see Section 10.1 in Kosorok (2008)).

4 A semi-parametric model of first-price auctions with independent private values without auxiliary data

A nonparametric approach in the k–rationalizability model is not available any more if there is no auxiliary data from ascending auctions. In other words, for any given k, any distribution of bids can be rationalized for some distribution of private values. So, without any assumptions on the distribution of private values k–rationalizability can not be rejected for any finite k. The following result proves this assertion. I omit conditioning on the observable characteristics of the object X for simplicity of notation. I consider a first-price sealed-bid auction with n symmetric bidders. Private values are independent and distributed according to cdf $F_0 : [0,1] \to [0,1]$.

Result 4.1. *Any distribution of bids $G(t)$ such that $g(t) = G'(t)$ exists and is positive and continuous can be rationalized for any given k by some distribution of private values F_0, such that $f_0(t) = F_0'(t)$ exists and is positive.*

Result 4.1[19] shows that if there is no auxiliary data from ascending or second-price auctions, the only way to test k–rationalizability is to assume some parametric form for the distribution of private values.

Assume that the distribution of private value F_0 is a member of a parametric family $\mathcal{F} = \{F_\theta : \theta \in \Theta\}$, such that all F_θ have the same support $[0,1]$ (for simplicity). In the IPV model, as I showed in the extension section, we can focus on a lower-envelope for this family, $\underline{F}_0(v) \equiv \inf_{F\in\mathcal{F}} F(v)$. This lower envelope can be used to construct an upper envelope for the bound $\overline{B}_k(v)$ (denoted as $\overline{\overline{B}}_k(v)$). The features of $\mathcal{F}$ will determine whether $\overline{\overline{B}}_k(v)$ will be nontrivial (i.e, whether it will be bounded away from the 45-degree line) and also whether this bound is sharp. Under technical conditions on the parametric family $\mathcal{F}$, $\underline{F}_0$ is a distribution function with density $\underline{f}_0$. In addition, $C_1 \le \underline{f}_0 \le C_2$ on its support $[0,1]$ (See Lemma 1 in Aryal et al. (2016)).

Remark 4.1. If the lower envelope $\underline{F}_0$ is a member of the parametric family $\mathcal{F}$, then the upper envelope $\overline{\overline{B}}_k$ will be a sharp bound for $\overline{B}_k$. More generally, this will be the case if, $\forall\, \epsilon > 0$, there exists $F \in \mathcal{F}$ such that $\underset{v\in[0,1]}{sup} \left|\overline{\overline{B}}_k(v) - \overline{B}_k^F(v)\right| < \epsilon$. Sharpness can result even if $\underline{F}_0 \notin \mathcal{F}$.

The following three examples illustrate the importance of the choice of the parametric family:

1. If $\mathcal{F} = \{F_\theta(t) = t^\theta$, s.t. $\theta > 0\}$, the lower envelope is $\underline{F}_0(t) = 0$ for $t \in [0,1)$ and $\underline{F}_0(1) = 1$. As I showed in the proof of Result 4.1, k–rationalizability can not be rejected for any finite k for this family.

2. If $\mathcal{F}$ is a family of truncated exponential distributions on $[0,1]$.

[19] Prove by induction of the Result 4.1 can be find in the supplementary material.

$$\mathcal{F} = \left\{ F_\theta(t) = \frac{e^{-\theta t} - 1}{e^{-\theta} - 1}, \text{ where } \theta > 0 \right\}.$$

In this case the lower envelope is $\underline{F}_0(t) = t$ (the $U[0,1]$ distribution). It is not a member of the family $\mathcal{F}$, but it is a limiting case when $\theta \to 0$, i.e. $\underline{F}_0(t) = \lim_{\theta \to 0} F_\theta(t)$, for all $t \in [0,1]$. Thus, the resulting bounds $\overline{\overline{B}}_k$ are sharp and informative. This example is illustrated in supplementary materials.

3. If $\mathcal{F}$ is a family of truncated normal distributions in $[0,1]$, with mean $\mu = 0.5$ and variance $\sigma^2 \geq \delta > 0$, then the lower envelope is

$$\underline{F}_0(t) = \begin{cases} F_\delta(t), & \text{if } t \in [0, 0.5], \\ t, & \text{if } t \in [0.5, 1]. \end{cases}$$

In this case the bounds are meaningful but not sharp.

If $\overline{\overline{B}}_k$ is not a sharp bound, or if we want to conduct inference both on θ and k, we can use the conditional moment inequalities implied by $k-$rationalizability and proceed as described in extensions for the parametric approach.

5 Empirical illustration: testing $k-$rationalizability in USFS timber auctions

As an illustration I apply my testing methodology to USFS timber auctions. The advantage of existing data is the ability to combine information from ascending and sealed-bid auctions of timber tracts in a way that is compatible with the assumptions of the nonparametric test developed in Section 3.2. Combining information from both auction formats in timber auctions has been done before, for example in Lu and Perrigne (2008) (to estimate risk aversion) and in Athey et al. (2011) (to study bidders participation and auction design). An overview of the data and the results of reduced-form tests for the presence of collusion (which is ruled out in the $k-$rationalizable model) of the type proposed in Porter and Zona (1993) are provided in the supplementary material. Based on those diagnostic tests, competitive bidding cannot be rejected at a 1% significance level.

Following Athey et al. (2011) I classify bidders into two types: "small" (s) and "large"(ℓ) according to the number of workers, with the threshold being 150 workers. To make things tractable, I focus on auctions with two bidders. In this case, there are three possible configurations of auctions: two symmetric and one mixed. I test for k separately for each one of the three possible auction configurations $((s,s),(s,\ell),(\ell,\ell))$. In the data, the proportion of observations according to configuration is the following (s,s) is 44%, (s,ℓ) is 28% and (ℓ,ℓ) is 27%. I allow for k to be potentially different for each type, and to also vary depending on the type of the opponents. This provides greater flexibility and a more realistic approximation of bidding behavior.

My results show that values of k as low as 2 can be rejected in asymmetric auctions, while bidding behavior is consistent with larger values of k in symmetric auctions. The results indicate that large bidders underestimate the degree of rationality of small bidders (as measured by $k_{s\ell}$). I perform counterfactual analysis to quantify the economic losses that large bidders incur as a result of their incorrect beliefs. Direct tests for BNE of the type described in Section 5.4 also reject this behavior, as I find evidence of *overbidding and underbidding* relative to BNE, a pattern that cannot be explained by risk aversion or by the usual models of collusion.

5.1 Data description

My data set includes all US Forest Service (USFS) timber sales in Region 6 (Oregon and Washington) in period 1994-2007. Timber tracts are sold by the USFS using both sealed-bid first-price and ascending auctions. My sample includes 3484 observations in total: 902 sealed-bid first-price auctions and 2359 ascending (second-price) auctions. For each auction I observe the format of the auction, the number of participants, their identity and characteristics such as number of workers and manufacturing class. For each auction there is a rich collection of characteristics of the timber tract: geographical characteristics (state, county, national forest, acres to be harvested), total volume of timber in thousand of board feet (mbf), species (I calculate Herfindahl index for the concentration of species of the tract to take into account the diversity of the species), road construction costs, logging costs, manufacturing costs, total costs, the USFS advertised value (reserve price), the date of the auction and the contract length. In first-price auctions I observe all bids submitted, and in ascending auctions I observe the highest bid of each participant. However, the test is carried out by focusing on transaction price (highest bid) in the ascending auctions, as the econometric test in Section 3.2 describes. Summary statistics for all auctions and for the subpopulation of ascending and sealed-bid auctions are given in the supplementary material.

My test allows the observable characteristics X to have different distributions in different types of auctions, as long as their support has a nonempty intersection. There appear to be some systematic differences between ascending and sealed-bid auctions. The reserve price per mbf is higher in sealed-bid auctions than in ascending auctions (\$188.21 vs. \$152.06), but the standard deviation is large in both cases. A similar pattern is observed for transaction price, which is highly correlated (0.71) with the advertised price. The volume of timber is on average two and a half times larger in ascending auctions. The average Herfindahl index is similar in both types of auctions. The average of 0.75 indicates that the timber tracts are homogeneous in terms of species, which makes it easier to compare different sales. There is no obvious difference in competition (based on the number of participants), since the average number of participants are equal in both types of auctions. Costs of road construction, logging and manufacturing are higher in ascending auctions. To summarize, the observable sale characteristics and auction outcomes differ in a number of variables that have large variation in both types. This discrepancy is perfectly compatible with the assumptions of my econometric test as long as the support of the conditioning variables chosen for X has a nonempty intersection between both auction formats.

It is also worthwhile to study the question of what affects the USFS choice of the auction type. To address this question, I consider a logit regression where the dependent variable equals one for the sealed-bid auctions and zero for the ascending auctions. Among the explanatory variables I include: reserve price, volume, three types of costs and the Herfindahl index. I also control for the state variable and include dummies for years and quarters of the sale. The results of the logit regression are represented in the supplementary material. Sealed-bid auctions appear to be chosen more frequently for tracts with higher appraisal value (per unit of timber), where the distribution of species is relatively more homogeneous. Ascending auctions appear to be chosen more frequently for large-volume tracts. Nonrandom assignment of an auction format is entirely consistent with the assumptions of my test, which only require that, for a given (fixed) vector of tract observable characteristics X, the distribution of bidders' values is not affected by the auction format (i.e, $F_0(\cdot|X)$ is the same whether the tract is sold through an ascending or a sealed-bid auction).

The total number of unique bidders is 932. Out of these, 454 participated in at least two auctions. If I split these figures across auction formats, 551 unique bidders participated in first-price auctions and 623 in ascending auctions. A total of 242 bidders participated in both auction formats. This represents approximately 25% of all bidders in the sample, but 53% of those who participated in at least two auctions. It is important to compare the characteristics of bidders across both auction formats. p-value for the test of equality of the distributions of bidders size in two auction formats is 0.21. Thus, there is no evidance of sorting of bidder according to the auction format.

5.1.1 Selection of X, the vector of conditioning variables

As I mentioned above, the test focuses on auctions with two bidders, so the analysis is conducted conditional on $n = 2$. More precisely, I condition on the specific configuration of types ((s, s),

(ℓ, s), (ℓ, ℓ)). However, I still need to consider which additional conditioning variables to include in X. Ideally, all relevant observable auction characteristics would be included in X; however because my test is nonparametric this would create curse of dimensionality in a relatively small sample size like mine. For this reason I have to choose X carefully. A reduced-form exploratory analysis of the determinants of the transaction price per volume in auctions with two players reveals that the only auction characteristic that has explanatory power in both types of auctions is the reserve price per volume. For this reason and to mitigate the effect of the curse of dimensionality, *reserve price is the only conditioning variable I include in* X[20].

My test requires that the supports of the corresponding distributions of the reserve price have a nonempty intersection. In fact, while the distributions are not necessarily the same (they are not assumed to be), the supports appear to be the same. Moreover, Kolmogorov-Smirnov test for the equality of the distributions has a p-value $= 0.0389$. The validity of that assumption appears to be supported by the data in my empirical illustration.

5.2 Collusion

Patterns of bidding behavior consistent with collusion in ascending timber auctions in the Pacific Northwest (Region 6) have been econometrically documented, for the period 1975-1981, by Baldwin et al. (1997). Even though this is the same region in my sample, those results correspond to a much earlier time period. Baldwin et al. (1997) document significant changes to the USFS sales program after 1982. Notably, a shortening of the time allowed to complete the clear cut of the tract as well as a substantial increase in bid bonds required from winners at the time of the sale. These changes can be argued to have fundamentally changed the incentives for collusion. Nevertheless, the methodology and approach in this paper can be robust to the presence of collusion in the auxiliary sample of ascending auctions. If all drop-out bids are observed in the sample of ascending auctions, the distribution of bidders' values can still be nonparametrically identified even in the presence of collusive behavior as long as collusion is efficient (i.e, the cartel leader is the bidder with the highest valuation and the researcher knows at least one competitive bidder) as in Kaplan et al. (2017).

Since, in my empirical application, I focus only on auctions with two players, the questions about collusion in ascending auctions is easy to address. Assume there is a set of colluding bidders. Note, that in contrast to Kaplan et al. (2017) and Schurter (2017), in my case the set of bidders in an auction is not fixed, i.e. different players may participate in a two bidders

[20] The dependent variable is the transaction price. The regression was estimated over the subsample of auctions that were run and had the transaction price reported. The estimation method is Nonlinear Least Squares with a sample-selection correction term that assumes a mean-zero normally distributed error term (whose variance is estimated along with the parameters reported)

auction. So, there are three possible cases: both bidders are competing; one is competing, the other is colluding (Since there are only two bidders in an auction and only one member of cartel decided to participate, there is no one to collude with! In this case, both players compete and the transaction price equals the second highest private value); both bidders are colluding (Since there is no one to compete with, the transaction price equals the reserve price).

I exclude those ascending auctions where the transaction price is only 1% higher than the reserve price. Those auctions are the most suspicious for collusion and may effect the consistent estimator of the distribution of private values. The remaining observations in the ascending auctions are consistent with competitive bidding and using this data I consistently estimate the distribution of private values.

Collusion in first-price auctions is less likely than in the ascending auctions because bidding ring is not self-enforcing in this case. However, the $k-$rationalizable bounds in first-price auctions presuppose non-cooperative bidding behavior in those auctions. Therefore, it is worthwhile exploring whether there is significant evidence of collusion in the sample of first-price auctionsIn Supplementary materials, I carry out the type of reduced-form exploratory tests for collusion proposed by Porter and Zona (1993). The analysis splits all bidders into two groups: competitive and potential colluders and specifies a functional form for bidding functions. Under the null hypothesis of competitive behavior, the features of the bidding functions should be the same for both groups. Tests are carried out for bid-levels as well as bid-rankings (to detect the presence of "phantom bids"). After running several tests for different subsets of potential colluders, my results cannot reject the hypothesis of competitive bidding at a 1% significance level.

5.3 Implementation of the test

I apply the test described in Section 3.2 for auctions with $n = 2$, conditional on X =reserve price (per unit of timber) and conditional on the configuration of bidder types observed $((s, s), (s, \ell), (\ell, \ell))$. This allows me to test separately four potentially distinct values of $k_{ij}, i \in \{s, \ell\}, j \in \{s, \ell\}$ the number of iterated steps a firm of type i performs to submit its bid against a firm of type ℓ.
I treat the reserve price as a continuous variable. So, the number of continous conditioning variables is $c = 1$ and the order of bias-reducing kernel must be at least $M = 2c + 1 = 3$. I use bias-reducing kernel of order 3, symmetric around zero with bounded support $[-1, 1]$.

$$k(u) = \left(C_1(1 - u^2)^2 + C_2(1 - u^2)^4\right) 1(|u| \leq 1),$$

where C_1 and C_2 are chosen in such way that $\int_{-1}^{1} k(u)du = 1$ and $\int_{-1}^{1} u^2 k(u)du = 0$ for $j = 2$. Symmetry around zero ensures that $\int_{-1}^{1} u^j k(u)du = 0$ for any odd j. The bandwidths used in

Table 1: Results from the tests for $k-$rationalizability

	k_{ss}		$k_{s\ell}$		$k_{\ell\ell}$		$k_{\ell s}$	
	$\widehat{t}_L$	p-value	$\widehat{t}_L$	p-value	$\widehat{t}_L$	p-value	$\widehat{t}_L$	p-value
$k=2$	0.009	0.496	1.118	0.132	0.503	0.307	2.108	0.018
$k=3$	0.118	0.453	2.060	0.020	0.904	0.183	2.326	0.010
$k=4$	0.158	0.437	2.095	0.018	1.099	0.136	2.354	0.009
$k=5$	0.162	0.434	2.096	0.018	1.134	0.129	2.358	0.009
$k=6$	0.166	0.434	2.096	0.018	1.138	0.128	2.358	0.009

the nonparametric estimators are normalized to $0.5 \cdot \widehat{\sigma}(\cdot)$ for reserve price X, for bids b and for transaction price P respectively, where $\widehat{\sigma}(\cdot)$ is a standard deviation in the corresponding sample. The tuning parameters γ_{L_l} and γ_{L_2} are set to be 0.004. The tuning parameter λ_L is set below machine precision.

The test is constructed for a fixed value of k. A naïve way to find the largest value of k consistent with data would be to start with $k=2$ and increase k till the first rejection. This approach has two problems: first, if bidders follow a rationalizable strategy, the process never stops; second, even in the process stops at some finite k, the procedure can have size distortion because of the sequential testing. While $\widehat{\mathcal{T}}_k$ is monotone with respect to k, monotonicity of $\widehat{t}_k$ is unclear. Thus, I proceed in the following way: first, I calculate $\widehat{\mathcal{T}}_k$ starting from $k=2$ and increasing k. Since upper bounds $\overline{B}_k$ monotonically converge and calculation is numerical, $\widehat{\mathcal{T}}_k$ converges fast. In my data set, $\widehat{\mathcal{T}}_k$ converges in $k=6$ steps for all configurations of bidder types. I denote this k as k_{max}[21]. Next, the variance σ^2_{k,L_1,L_2} described in Equation (3.9) is estimated by bootstrap with 1000 bootstrap draws for $k=2,\ldots,k_{max}$ and $\widehat{t}_k = \widehat{\mathcal{T}}_k / \sqrt{\max_{k=2,\ldots,k_{max}}\{\lambda_L/L, \widehat{\sigma}^2_{k,L_1,L_2}\}}$ to insure monotonicity of the corresponding p-values. Table 1 summarizes the results of the test. At a 5% significance level, $k_{s\ell} \geq 3$ is rejected and $k_{\ell s} \geq 2$ is also rejected, while I cannot reject any $k \leq k_{max}$ for k_{ss} or $k_{\ell\ell}$. The results appear to indicate a fundamental difference in beliefs when bidders face an opponent of their type compared to the case when they face an opponent of a different type. Furthermore, there is a misalignment in beliefs across types: when they face each other, large bidders think that small ones bid above the $k=2$ rationalizable bounds with positive probability; however, my results indicate that this is not true, since $k=2$ was not rejected for small bidders. A counterfactual exercise below attempts to quantify the monetary loss in expected payoff that results from large bidders' incorrect beliefs.

[21] A theoretical result for the choice of k_{max} and the rate of convergence of $\widehat{\mathcal{T}}_k$ is beyond the scope of this paper and left for the future research.

5.4 Tests for risk-neutral BNE

The presence of auxiliary data from ascending auctions facilitates the construction of a direct test for risk-neutral BNE. Moreover, I can test the form in which risk-neutral BNE is violated by performing inequality-tests to see if bidders bid above or below BNE (or both). This can help to rule out, for example, risk aversion as the source behind those violations since risk-averse BNE bids cannot be below risk-neutral BNE bids in my setting (see Section 4.1 in Krishna (2010)). On the other hand, $k-$rationalizability is consistent with both types of violations of risk-neutral BNE. For simplicity, my BNE tests focus on auctions where both bidders are of the same type. Those are the auctions where no value of k between 2 and 6 could be rejected; therefore it is relevant to investigate if these bidders are playing a risk-neutral BNE. Symmetry yields a straightforward closed-form expression for the BNE bidding function (see Section 2.3 in Krishna (2010)). These are given by

$$b_{BNE}^{\imath\imath}(v|x) = \frac{1}{F_{0,\imath}(v|x)} \int_{\underline{v}}^{v} \frac{t \cdot f_{2\imath,P|X}(t|x)}{2 \cdot (1 - F_{0,\imath}(t|x))} dt, \quad \text{for } \imath = s, \ell.$$

where, as before, ss and $\ell\ell$ refer to two-bidder auctions involving two small and two large bidders, respectively. These BNE bidding functions are strictly increasing and invertible in v conditional on x and $v_{BNE}^{\imath\imath}(\cdot|x), \imath = s, \ell$ their respective inverse functions. Let $G_1^{\imath\imath}(\cdot|x), \imath = s, \ll$ denote the conditional cdf of bids in two-bidder first-price auctions given x. For a given $t \in (\underline{b}, \overline{b})$ denote

$$\phi_{BNE}^{\imath\imath}(t|x) = F_{0,\imath}\left(v_{BNE}^{\imath\imath}(t|x)|x\right) - G_1^{\imath\imath}(t|x), \quad \text{for } \imath = s, \ell.$$

One-sided violations of risk-neutral BNE can be tested via stochastic dominance restrictions, using the above functionals.

5.4.1 A test for no overbidding above risk-neutral BNE

I test the null hypothesis that bidders never bid above the risk-neutral BNE in two-player auctions involving bidders of the same type and do this separately for small and large bidders. Framing it as a stochastic-dominance restriction, this is a test of the null hypotheses

$$H_0 : \phi_{BNE}^{\imath\imath}(t|x) \leq 0, \quad \text{for all } x \in \mathscr{X}, t \in \mathscr{B}(x) \text{ and } \imath = s, \ell \text{ separately.}$$

The conditions leading to Theorem 3.1 produce an analogous result in this case. A test can be constructed in the same way as the test for $k-$rationalizability. Rejecting H_0 indicates the presence of bids above risk-neutral BNE, which indicates a violation of BNE that cannot be explained by the usual models of collusive behavior that predict bids bounded above by competitive (i.e, BNE).

5.4.2 A test for risk-averse BNE

Testing the null hypothesis that bidders never bid below the risk-neutral BNE can be framed as a test for

$$H_0 : \phi^{\imath\imath}_{BNE}(t|x) \geq 0, \quad \text{for all } x \in \mathscr{X}, t \in \mathscr{B}(x) \text{ and } \imath = s, \ell \text{ separately.}$$

That is, the reverse inequality as above. Rejecting H_0 indicates the presence of bids below BNE, which rejects equilibrium behavior but it also rejects risk-averse BNE as the true model.

5.4.3 Two-sided BNE test

Testing the null hypothesis that all bidders play according to the risk-neutral BNE is equal to testing

$$H_0 : \phi^{\imath\imath}_{BNE}(t|x) = 0, \quad \text{for all } x \in \mathscr{X}, t \in \mathscr{B}(x) \text{ and } \imath = s, \ell \text{ separately.}$$

In this case, I use similar statistics. The only difference is in the step of constructing $\Lambda^{\imath\imath}_{BNE}(x), \imath = s, \ell$ in (3.5):

$$\Lambda^{\ell\ell}_{BNE}(x) = \int_{t \in \mathscr{B}(X)} \left[\left(\phi^{\ell\ell}_{BNE}\right)_+ + \left(-\phi^{\ell\ell}_{BNE}\right)_+ \right] dQ(t|x), \quad \text{for } \imath = s, \ell.$$

5.4.4 Results for BNE tests

I implement the BNE tests described previously using the same kernels and bandwidths described above for the $k-$rationalizability tests. Results are shown in Table 2

Table 2: Results of BNE tests in auctions with bidders of the same type.

	Small firms		Large firms	
$\boldsymbol{H_0}$	statistic	p-value	statistic	p-value
No overbidding above risk-neutral BNE	1.56	0.059	2.39	0.008
Risk-averse BNE	2.23	0.013	0.77	0.221
Risk-neutral BNE	2.41	0.008	2.39	0.008

At a significance level of 5%, the results in Table 2 reject risk-neutral BNE in all kinds of two-bidder symmetric auctions. However, they suggest that violations to equilibrium behavior are different in nature across the two types of bidders. Small bidders appear to depart from equilibrium behavior by submitting bids below risk-neutral BNE, while large bidders overbid BNE. Thus, the inability to reject $k-$rationalizability in symmetric auctions for the range $[2, 6]$

is not owed to the presence of BNE behavior in those auctions. However, I cannot rule out that bids are rationalizable (i.e, consistent with $k \to \infty$) in those types of auctions, since the estimated bounds in the data are indistinguishable for $k \geq 6$.

5.5 A counterfactual exercise to estimate the economic losses from incorrect beliefs

The results from the $k-$rationalizability tests in Table 1 describe a misalignment in beliefs. In asymmetric auctions with two participants, large bidders appear to believe that small ones violate the $k = 2$ bounds with positive probability; but our results indicate that small bidders in fact bid below the $k = 2$ bounds with probability one. Therefore, *large bidders are using conjectures that are too pessimistic given the actual bidding strategies of small bidders*. This can lead to losses in expected payoff and overbidding for large bidders in those auctions. In this section I perform counterfactual analysis to estimate these effects.

To perform this exercise I construct an estimator for large bidders' current bidding functions under the assumption that, conditional on the realization of X, bidding functions are strictly increasing transformations of bidders' values (BNE bidding functions is a special case). Let

$b_{\ell}^{\ell,s}(\cdot|x)$ =the bidding functions of large firms in two-bidder asymmetric auctions.

$G_{1,\imath}^{\imath,s}(\cdot|x)$ =cdf of bids for bidders of type $\imath$ in two-bidder asymmetric first-price auctions, $\imath = s, \ell$.

$F_{0,\imath}(\cdot|x)$ =cdf of values for bidders of type $\imath$ (assumed to be the same in first-price and in ascending auctions), $\imath = s, \ell$

For a given t, under the invertibility assumption described above these functions can be nonparametrically estimated as the solution to the following equation,

$$\widehat{G}_{1,\ell}^{\ell,s}(t|x) = \widehat{F}_{0,\ell}\left(\widehat{b}_{\ell}^{\ell,s^{-1}}(t|x)\big|x\right) \tag{5.1}$$

5.5.1 Three measures of expected payoffs

To analyze the impact on expected payoffs I focus on three measures of expected payoffs (as functions of values v and reserve price x) for large bidders:

Benchmark case. This is intended to be a measure of large bidders' current expected payoffs given the bidding strategies estimated in (5.1). It is a measure of the status-quo given by

$$\pi_{benchmark}^{\ell}(v|x) = (v - b_{\ell}^{\ell,s}(v|x)) \cdot \boldsymbol{P}\big(b_s^{\ell,s} \leq b_{\ell}^{\ell,s}(v|x)\big|x\big) = (v - b_{\ell}^{\ell,s}(v|x)) \cdot G_{1,s}^{\ell,s}\big(b_{\ell}^{\ell,s}(v|x)\big|x\big).$$

Benchmark expected payoffs for large bidders are therefore estimated as

$$\widehat{\pi}^{\ell}_{benchmark}(v|x) = (v - \widehat{b}^{\ell,s}_{\ell}(v|x)) \cdot \widehat{G}^{\ell,s}_{1,s}\big(\widehat{b}^{\ell,s}_{\ell}(v|x)\big|x\big).$$

Boundary case. Given the results in Table 1, the most conservative assessment of expected payoffs for large bidders that is consistent with small bidders' *actual behavior* corresponds to the conjecture that small bidders bid at the upper bound for $k = 2$ rationalizable bids. Let $\overline{B}^{\ell,s}_{s,2}(\cdot|x)$ denote these bounds and let $\overline{v}^{\ell,s}_{s,2}(\cdot|x)$ denote its inverse function. The "boundary case" measures large bidders' *optimal expected payoff* for the most pessimistic conjectures that are consistent with the results. Boundary expected payoffs are therefore given by

$$\pi^{\ell}_{boundary}(v|x) = \underset{b \geq \underline{v}}{max}\left\{(v-b) \cdot \boldsymbol{P}\big(\overline{B}^{\ell,s}_{s,2}(v_s|x) \leq b\big|x\big)\right\} = \underset{b \geq \underline{v}}{max}\left\{(v-b) \cdot F_{0,s}\big(\overline{v}^{\ell,s}_{s,2}(b|x)\big|x\big)\right\}$$

They can be estimated as

$$\widehat{\pi}^{\ell}_{boundary}(v|x) = \underset{b \geq \underline{v}}{max}\left\{(v-b) \cdot \widehat{F}_{0,s}\big(\widehat{\overline{v}}^{\ell,s}_{s,2}(b|x)\big|x\big)\right\}.$$

Optimal case. Finally, I consider the case where large bidders best-respond to the current bidding strategies of small bidders, as indicated by their bidding functions $b^{\ell,s}_{s}(\cdot|x)$. This measure of expected payoffs is given by

$$\pi^{\ell}_{optimal}(v|x) = \underset{b \geq \underline{v}}{max}\left\{(v-b) \cdot G^{\ell,s}_{1,s}(b|x)\right\}$$

These payoffs can be estimated as

$$\widehat{\pi}^{\ell}_{optimal}(v|x) = \underset{b \geq \underline{v}}{max}\left\{(v-b) \cdot \widehat{G}^{\ell,s}_{1,s}(b|x)\right\}$$

The goal in this counterfactual exercise is to estimate the economic impact for large bidders of having incorrect beliefs about small bidders.

5.5.2 Results from comparing current expected payoffs against counterfactuals

Current bidding strategies by large firms can lead to substantial losses in expected payoff for a wide range of reserve prices. A simple switch to the boundary case –which uses the most pessimistic assessment consistent with the test results– can lead to substantial improvements. Naturally, optimal expected payoffs dominate both the benchmark and boundary cases, but the latter is very close to optimal for a wide range of reserve price values in the data. For some values of the reserve price expected payoff in the benchmark scenario is higher than in the boundary

Table 3: Counterfactual results: Difference in large bidders' expected payoffs (in 2010 real dollars).

	Difference in expected profits per-unit of timber			Difference in expected profits by total volume of timber		
	Boundary minus benchmark	Optimal minus benchmark	Optimal minus boundary	Boundary minus benchmark	Optimal minus benchmark	Optimal minus boundary
All auctions with two asymmetric bidders						
Median	\$692	\$1,728	\$506	\$302,973	\$609,928	\$379,682
75th percentile	\$6,919	\$8,377	\$2,167	\$5,929,771	\$7,629,625	\$1,538,886
Auctions with two asymmetric bidders won by large bidders						
Median	\$324	\$1,726	\$463	\$152,177	\$501,003	\$234,728
75th percentile	\$7,089	\$7,681	\$1,906	\$2,109,349	\$2,302,707	\$1,174,922

Note: The median, 75th and 90th percentiles for advertised value (in total volume) in this sample are \$81,325, \$282,961 and \$781,137, respectively.

scenario which indicates that for these values of the reserve price current beliefs of large firms is closer to the true bidding behavior of small firms that the most pessimistic one. The model makes no predictions about how bids, value distributions and, therefore, expected payoffs, shift with x (reserve price).

I compare the difference in expected payoffs, conditional on reserve price, at the level of each individual auction. Tables 3-5 describe the median and the 75^{th} quantiles, for the auctions in the sample. They include results for all auctions with two asymmetric bidders, and for the specific auctions that were won by large bidders.

The main findings can be summarized as follows.

1. The median improvement in expected payoff from correcting beliefs towards the most pessimistic case consistent with the data raises expected payoffs in approximately 5%. This amounts to a median gain per auction of approximately $\$300K$ (2010 dollars), which corresponds to approximately 3.75 times the median advertised value (reserve price) of these tracts.

2. A simple improvement in beliefs towards the boundary case can lead to substantial economic improvements in expected payoff. For 25% of the auctions in the sample, the improvement is at least six million 2010 dollars, which corresponds to approximately a 260% improvement relative to the current expected payoff.

3. Gains in expected payoff from using boundary beliefs for the subsample of auctions that were won by large bidders are comparatively smaller, but still substantial: the median gain is about 2% (approximately 2 times the median advertised value) but the 75^{th} percentile corresponds to an improvement of about 160%.

4. Responding optimally to small bidders' actual bidding strategies is naturally the best case scenario, but the improvements over the boundary-beliefs case are relatively minor (between

Table 4: Counterfactual results: Percentage difference in large bidders' expected payoffs.

All auctions with two asymmetric bidders			
	$\frac{\text{Boundary}}{\text{Benchmark}}$ (%)	$\frac{\text{Optimal}}{\text{Benchmark}}$ (%)	$\frac{\text{Optimal}}{\text{Boundary}}$ (%)
Median	104%	112%	108%
75th percentile	258%	265%	124%
Auctions with two asymmetric bidders won by large bidders			
	$\frac{\text{Boundary}}{\text{Benchmark}}$ (%)	$\frac{\text{Optimal}}{\text{Benchmark}}$ (%)	$\frac{\text{Optimal}}{\text{Boundary}}$ (%)
Median	102%	105%	107%
75th percentile	150%	157%	122%

Table 5: Counterfactual results: Evidence of overbidding. Summary statistics for $\frac{b^{\ell}_{benchmark}(\cdot)}{b^{\ell}_{boundary}(\cdot)}$ (in % terms) in auctions won by large bidders.

	All auctions	High-value auctions		
		Auctions with reserve price in upper 25th percentile	Auctions with reserve price in upper 10th percentile	Auction with highest reserve price in the sample
Median	102%	113%	141%	145%
75th percentile	127%	142%	145%	149%

20%-25%). This suggests that small bidders' actual bidding strategies are not very far from the $k = 2$ bounds when they play against large competitors.

5. Large bidders' incorrect beliefs lead to overbidding, and the proportion of overbidding is substantially larger in high-value auctions. Table 5 focuses on auctions won by large bidders (where they actually had to pay the bids they submitted). Let us focus on the 25% most valuable tracts. Comparing the observed bids against the counterfactual optimal bids in the boundary case, the median proportional amount of overbidding is 13% and the 75^{th} percentile is 42%. Moving to the 10% most valuable tracts, these proportions are 42% and 45%, respectively.

5.6 Robustness check

I check how sensitive my results to the choice of tunning parameters. The magnitude of $\phi(t|x)$ is bounded between -1 and 1. Results do not change at 5% significance level if I halve or double the tunning parameters γ_{L_1} and γ_{L_2}. Results change slightly (see Table 6) for $\gamma_{L_1} = \gamma_{L_2} = 0.001$, i.e. 4 times smaller than in the main specification.

5.6.1 Linear index

Ideally, I would like to use all observable characteristics of the auction as conditional variables. Since, the approach is nonparametric, I have a "curse of dimantionality" even with a small number of continuous conditioning variables. In my data set, advertising price is a natural candidate for

Table 6: Results from the tests for $k-$rationalizability conditioning only on the reserve price ($\gamma = 0.001$)

	k_{ss}		$k_{s\ell}$		$k_{\ell\ell}$		$k_{\ell s}$	
	$\widehat{t_L}$	p-value	$\widehat{t_L}$	p-value	$\widehat{t_L}$	p-value	$\widehat{t_L}$	p-value
$k=2$	0.107	0.458	1.116	0.132	0.522	0.301	2.149	0.016
$k=3$	0.216	0.415	1.548	0.061	0.924	0.178	2.366	0.009
$k=4$	0.256	0.399	1.572	0.058	1.120	0.131	2.395	0.008
$k=5$	0.263	0.396	1.573	0.058	1.154	0.124	2.400	0.008
$k=6$	0.264	0.396	1.573	0.058	1.159	0.123	2.400	0.008

the unique conditioning variable that contains all needed information about the timber tract for sale. But in other application, the choice of the conditioning variable is not as straight forward. Thus, I explain how to use estimated linear index as a conditioning variable. The detail analysis of this extension is the supplemental materials.

Tables 7 and 8 represent results of testing k-rationalizability conditioning on the estimated linear index. Linear index is estimated with Tobit 1 model of the transaction price in both first price and ascending auctions. I take into account those auctions where no bidders decided to participate in an auction. Thus, for those auctions I do not observe the transaction price. I include advertising price, HHI of the species, roads, loggind and manufacturing costs, state, size of the tract, volume, contract duration and different time dummy variables. More results can be found in the supplementary materials.

Table 7: Results from the tests for $k-$rationalizability conditioning on the linear index ($\gamma = 0.004$)

	k_{ss}		$k_{s\ell}$		$k_{\ell\ell}$		$k_{\ell s}$	
	$\widehat{t_L}$	p-value	$\widehat{t_L}$	p-value	$\widehat{t_L}$	p-value	$\widehat{t_L}$	p-value
$k=2$	0.412	0.340	0.724	0.235	0.303	0.381	1.606	0.054
$k=3$	0.582	0.280	1.476	0.070	0.948	0.172	1.950	0.026
$k=4$	0.640	0.261	1.574	0.058	1.147	0.126	2.506	0.006
$k=5$	0.654	0.257	1.583	0.057	1.191	0.117	2.661	0.004
$k=6$	0.656	0.256	1.584	0.057	1.197	0.116	2.687	0.004

Table 8: Results from the tests for $k-$rationalizability conditioning on the linear index ($\gamma = 0.001$)

	k_{ss}		$k_{s\ell}$		$k_{\ell\ell}$		$k_{\ell s}$	
	$\widehat{t_L}$	p-value	$\widehat{t_L}$	p-value	$\widehat{t_L}$	p-value	$\widehat{t_L}$	p-value
$k=2$	0.478	0.316	0.787	0.216	0.322	0.374	1.631	0.051
$k=3$	0.649	0.258	1.542	0.062	0.967	0.167	1.974	0.024
$k=4$	0.707	0.240	1.641	0.050	1.166	0.122	2.530	0.006
$k=5$	0.721	0.236	1.651	0.049	1.210	0.113	2.687	0.004
$k=6$	0.724	0.235	1.651	0.049	1.217	0.112	2.707	0.003

6 Extensions

6.1 A nonparametric test when the distribution of private values is only partially identified.

A test for $k-$rationalizability can still be constructed if F_0 is not identified as long as a lower bound $\underline{F}_0$ for it can be identified. The resulting test would be less powerful but it would be robust to a much wider range of formats of ascending auctions. Suppose $\underline{F}_0(\cdot)$ is a lower bound for $F_0(\cdot)$, so that $\underline{F}_0(\cdot|x) \leq F_0(\cdot|x)$ for a.e x. Suppose F_0 is not identified but $\underline{F}_0(\cdot)$ is. This will be enough to identify an upper bound for $\overline{B}_k$, the $k-$rationalizable bounds, which I denote as $\overline{\overline{B}}_k(\cdot|x)$.

$$\overline{\overline{B}}_1(\cdot|x), \qquad \overline{\overline{B}}_k(v|x) = v - \max_{b \geq 0}(v-b)\underline{F}_0^{n-1}(\overline{\overline{B}}_{k-1}^{-1}(b|x)|x), k = 2, \cdot.$$

If $\underline{F}_0$ is a cdf for some distribution and pdf $\underline{f}_0$ exists and positive, the upper bounds $\overline{\overline{B}}_k(\cdot|x)$ remain strictly increasing functions that monotonically converge to some limit. Denote the inverse of $\overline{\overline{B}}_k(\cdot|x)$ by $\overline{\overline{v}}_k(\cdot|x)$. Then the stochastic dominance condition in (3.4) holds *only if*,

$$\underline{F}_0(\overline{\overline{v}}_k(t|x)|x) \leq G_1(t|x) \quad \forall\ t \in [\underline{b}, \overline{b}], x \in S_{2,X} \tag{6.1}$$

Given my data assumptions, a nonparametric lower bound for $F_0(\cdot)$ can be obtained using the results from Haile and Tamer (2003) that allows jump bidding, absence of bidding and the possibility of transaction price being smaller than the second highest bidder valuation. It also allows for a nonnegligible required bid-increment. All details are in supplementary material. In this case the test would be based on

$$\underline{\phi}_k(t|x) = \underline{F}_0\left(\overline{\overline{v}}_k(t|x)|x\right) - G_1(t|x)$$

instead of $\phi_k(t|x) = F_0(v_k(t|x)|x) - G_1(t|x)$. It has less power because $\underline{\phi}_k(t|x)$ only detects larger violations of $k-$rationalizability, but the results are robust to a wider range of behavior in the population of ascending auctions.

6.2 k-rationalizability with ambiguity

The assumption that the true distribution of private values is known by bidders can be replaced by the assumption of common prior set $\mathcal{F}$.

Assumption 6.1. (i) the prior set $\mathcal{F}$ is the same for all bidders (ii) true distribution of private values F_0 is a member of the common prior set $\mathcal{F}$

Denote $\underline{F}_0(t) \equiv inf_{f \in \mathcal{F}} F(t)$.[22] Now, there are two sources of ambiguity for bidders: the distribution of private values and beliefs about opponents' strategies. The following three ways of solving ambiguity lead to different upper bounds for k–rationalizable bids or even to point prediction of bidding strategies. In what follows, assume $\underline{F}_0$ is identified but perhaps F_0 is not.

Assumption 6.2. Bidders solve ambiguity related to the distribution of private values by best-responding to the most pessimistic belief given by the lower envelope $\underline{F}_0$

If assumptions 6.1 and 6.2 are satisfied then bidder i solves the following maximization problem given beliefs μ_i (see p.6 in Aryal et al. (2016)):

$$\pi^*(v_i; \mu_i) = \max_{b \geq 0}(v_i - b)\mathbf{P}_{\underline{F}_0}(b_{-i} < b|\mu_i)$$

Under technical conditions:

$$\overline{B}_1(v) = v, \quad \overline{B}_k(v) = \quad v - \max_{b \geq 0}(v - b)\underline{F}_0^{n-1}(\overline{B}_{k-1}^{-1}(b)), k = 2, \ldots. \tag{6.2}$$

Remark 6.1. In contrast to the semi-parametric approach in section 4, with ambiguity under Assumption 6.2, $\overline{B}_k(\cdot)$ is always a sharp upper bound for k–rationalizable bids.

The main first-order stochastic dominance inequality has the form:

$$F_0(\overline{B}_k^{-1}(t)) \leq G(t), \quad \forall t \in [\underline{b}, \overline{b}].$$

If F_0 is not identified[23] , the only testable implication is:

$$\underline{F}_0(\overline{B}_k^{-1}(t)) \leq G(t), \quad \forall t \in [\underline{b}, \overline{b}].$$

Assumption 6.2′. Bidders solve ambiguity about opponents' strategies by best-responding to the most pessimistic beliefs, but they do not necessarily know how other bidders solve this source of ambiguity.

If assumptions 6.1, 6.2 and 6.2′ are satisfied then for any fixed k players bid according to the bidding function:

$$\check{b}_k^*(v) = \text{argmax}_{b \geq 0}(v - b)\underline{F}_0^{n-1}(\overline{B}_{k-1}^{-1}(b)),$$

[22]One example of common prior could correspond to Haile and Tamer (2003) incomplete model of ascending auctions. In this case the true distribution of private values F_0 is not identified but $\mathcal{F}$ is. It is common for all bidders and the econometrician. Similar to semi-parametric approach in Section 4, it is enough to have common lower envelope $\underline{F}_0$.

[23]In contrast to Aryal et al. (2016) the distribution of private values F_0 and the lower envelope of the prior set $\underline{F}_0$ can not be identified from the sample of first-price auctions since I do not assume equilibrium behavior.

where $\overline{B}_k$ is defined by the iterative procedure (6.2).

In this case the model implies the following equality:

$$F_0(\check{b}_k^{*-1}(t)) = G(t), \quad \forall t \in [\underline{b}, \overline{b}].$$

If only the lower envelope $\underline{F}_0$ is identified, I can only test the inequality:[24]

$$\underline{F}_0(\check{b}_k^{*-1}(t)) \leq G(t), \quad \forall t \in [\underline{b}, \overline{b}].$$

Assumption 6.2′′. Bidders solve ambiguity about opponents' strategies by best-responding to the most pessimistic beliefs and they know that other bidders solve this source of ambiguity in the same way.

If assumptions 6.1, 6.2 and 6.2′′ are satisfied then for any fixed k players bid according the bidding function defined iteratively:

$$\check{\check{b}}_1^*(v) = v, \qquad \check{\check{b}}_k^*(v) = \quad \text{argmax}_{b \geq 0}(v-b)\underline{F}_0^{n-1}(\check{\check{b}}^{*}{}_{k-1}^{-1}(b)), \quad k = 2, \ldots .$$

Similar to the previous cases, the only testable implication is the inequality:[25]

$$\underline{F}_0(\check{\check{b}}_k^{*}{}^{-1}(t)) \leq G(t), \quad \forall t \in [\underline{b}, \overline{b}].$$

The implications in all three cases can be tested similarly to (3.4).

6.3 A semi-parametric model of first-price auctions with interdependent values and affiliated signals

Nonparametric analysis of first-price auctions with interdependent values has been done, for example, in Haile et al. (2004), Pinkse and Tan (2005) and Somaini (2015) (in affiliated values models), and in Hendricks and Porter (1988), Hendricks et al. (2003) and Li et al. (2000) (in common values models). These, and all existing *nonparametric identification results* in first-price auctions rely on the assumption that bidders use BNE bidding strategies, which could be an invalid in my model[26]. As a result, existing nonparametric identification results in first-price auctions cannot be applied here.

[24] If an upper bound $\overline{F}_0$ is also identified, as it is in Haile and Tamer (2003). Then another testable implication is $G(t) \leq \overline{F}_0(\check{b}_k^{*-1}(t)), \forall t \in [\underline{b}, \overline{b}]$. The test for the inequalities $\underline{F}_0(\check{b}_k^{*-1}(t)) \leq G(t) \leq \overline{F}_0(\check{b}_k^{*-1}(t)), \forall t \in [\underline{b}, \overline{b}]$ can be constructed similar to the one in Section 5.4.3.

[25] An upper bound can also be used as described in Footnote 24.

[26] Nonparametric identification results in first-price auctions have also been obtained in nonequilibrium models but under very specific assumptions concerning bidding behavior, such as cognitive-hierarchy or Level-k models (e.g, An (2017)). Once again, these models are too restrictive and only a special case of k-rationalizability.

Inference based on $k-$rationalizability can be carried out parametrically, which would be particularly useful in two cases:

(i) When there is no auxiliary data from ascending auctions.[27]

(ii) When I want to allow for interdependent values and affiliation.

Consider an auction with n bidders. Bidder i draws a signal s_i which is privately observed. Bidder i's valuation is given by a value function $v(s_i, s_{-i})$, which is strictly increasing in the first argument and assumed to be symmetric and nondecreasing in its last $n-1$ arguments. This function v is the same for all bidders. To simplify the exposition assume that the value function v does not depend on X. The joint distribution of signals is denoted as $F_s(s_1, \ldots, s_n)$, assumed to be symmetric in all its arguments. Let $F_0^s(\cdot)$ denote the marginal cdf of each signal (the same for all bidders given the above symmetry condition). B-S require signals to be *affiliated* (Milgrom and Weber (1982)), which is equivalent to the supermodularity of $\log f_s$. Beliefs and $k-$rationalizable bounds are now described conditional on the signal s_i observed by bidder i. Let X denote the vector of observable auction characteristics, and let $F_{s|X}(\cdot|X)$ and $F_0^s(\cdot|X)$ denote the joint cdf of signals conditional on X.

If I impose an additional restriction that bidders assume *increasing* bidding functions $b_{-i}(\cdot)$ of their opponents (a requirement that was not imposed with private values), the key result (winner's curse) is that $E\left[v(s_i, s_{-i})|s_i, b_{-i} \leq b\right] \leq E\left[v(s_i, s_{-i})|s_i\right]$ (the expected valuation conditional on the signal and the event of winning the object is bounded above by the expected valuation conditional on the signal only). Under these conditions, Theorem 6 in B-S shows that the $k-$rationalizable bounds are constructed in the same iterative way described in (3.3), replacing bidders' unobserved values with their conditional expectations given the signals. The iterative construction in (6.3) now becomes

$$
\begin{aligned}
&\overline{B}_1(s_i|X) = E\left[v(s_i, s_{-i})|s_i, X\right], \\
&\overline{B}_{k+1}(s_i|X) = E\left[v(s_i, s_{-i})|s_i, X\right] - \pi^*(s_i; \overline{B}_k(\cdot|X), X), \quad k = 1, 2, \ldots, \quad \text{where} \\
&\pi^*(s_i; \overline{B}_k(\cdot|X), X) = \\
&\underset{b \geq 0}{sup} \left\{ \left(\int_{\max_{j \neq i} \overline{B}_k(s_j|X) \leq b} \left(E[v(s_i, s_{-i}|s_i)] - b\right) dF_{s_{-i}|s_i, X}(s_{-i}|s_i, X) \right) \cdot \mathbf{P}\left(\max_{j \neq i} \overline{B}_k(s_j|X) \leq b \middle| s_i, X \right) \right\}
\end{aligned}
\tag{6.3}
$$

Consider a parametric model conforming to the above requirements, where both the value function and the distribution of signals (conditional on X) are parameterized, respectively, as $v(\cdot; \theta_1)$ and $F_{s|X}(\cdot|X, \theta_2)$. The vector of parameters is $\theta \equiv (\theta_1, \theta_2)$, which belongs in a parameter space

[27] Result 4.1 shows that it is impossible to reject the k–rationalizability model for any given k without auxiliary data and any assumptions on the form of the distribution of private values.

Θ that satisfies the symmetry and affiliation requirements described above. For a given θ, a parametric expression for $\overline{B}_k(s_i|X;\theta)$ for the corresponding bounds is given by (6.3). Under the affiliation, symmetry and monotone beliefs assumptions described above, the invertibility properties of the bounds (with respect to the signals s_i) are still satisfied (Theorem 12 in B-S).

Let $\overline{s}_k(\cdot|x,\theta)$ denote the inverse function of $\overline{B}_k(\cdot|x,\theta)$, and for a given $t \in \text{Supp}(S)$ define

$$\phi_k^s(t|x,\theta) = F_0^s\left(\overline{s}_k(t|x,\theta)|x,\theta_2\right) - G_1(t|x),$$

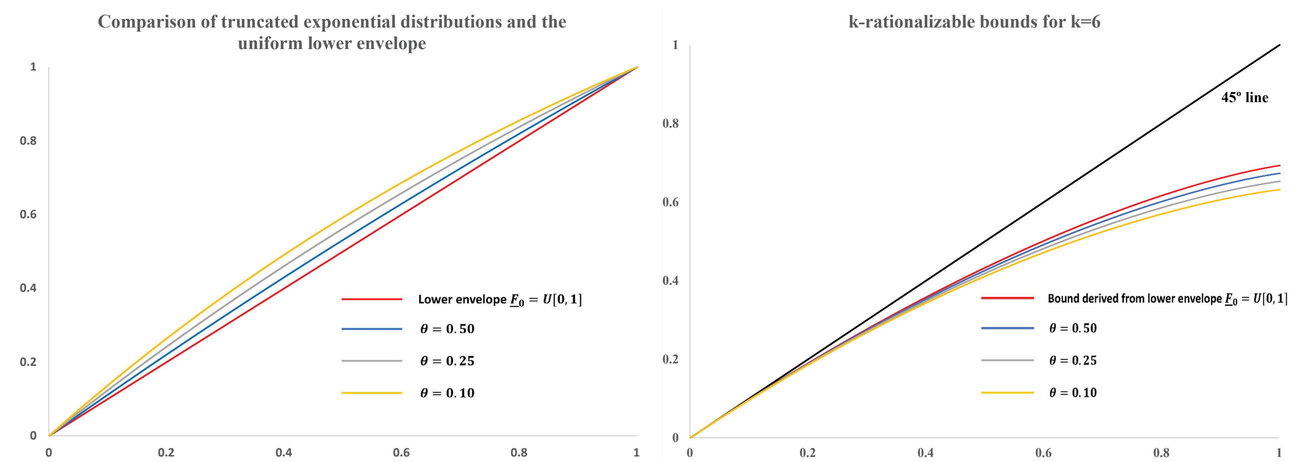
where, as before, $G_1(\cdot|X)$ represents the (nonparametrically specified) cdf of bids conditional on X. $k-$rationalizability requires that $\phi_k^s(t|x,\theta) \leq 0$ for all (t,x). Similar to (3.5), I can carry out inference based on these stochastic dominance conditions implied by the model.

Denote $\underline{\phi}_k^s(t|x) \equiv \inf_{\theta\in\Theta} F_0^s\left(\overline{s}_k(t|x,\theta)|x,\theta_2\right) - G_1(t|x)$. The following inequality:

$$\underline{\phi}_k^s(t|x) \leq 0, \quad \forall(t,x) \tag{6.4}$$

is a necessary condition for k–rationalizability, i.e. if there exists such $\widetilde{\theta} \in \Theta$ that $\phi_k^s(t|x,\widetilde{\theta}) \leq 0$ then the inequality (6.4) is satisfied. It can be tested similarly to (3.4)[28].

Figure 1: Truncated exponential family example. Distribution, lower envelope and $k-$rationalizable bounds (for $k = 6$).



[28]For a fixed k, the set of parameters Θ_I can be identified as in Aradillas-Lopez and Tamer (2008), where they also combine a parametric model with a nonparametric conditional functional through conditional inequalities. Corresponding confidence sets can be estimated using existing methods for conditional moment inequalities in Andrews and Shi (2011), Andrews and Shi (2013) or Chernozhukov et al. (2013). An empty confidence set corresponds to rejection of k.

6.4 Risk-aversion

In an IPV ascending auction with risk-averse bidders, bidding one's private value is still a weakly dominated strategy. Thus, the distribution of private values F_0 can still be identified and nonparametrically estimated from ascending auctions. However, identifying bidders' risk-aversion in a k–rationalizability model in first-price auctions is not possible even if auxiliary data from ascending auctions is available.

Risk aversion shifts the $k-$rationalizable bounds upwards. And if bidders are allowed to be arbitrarily risk-averse, these bounds can become arbitrarily close to the 45-degree line, making it impossible to test or reject any $k \geq 2$. In this case, the only fact that can be tested is whether players bid lower than their private value or not. This can be illustrated in the following example. Consider a constant relative risk-averse utility function:

$$u(x) = x^{\alpha}, \quad \alpha \in (0, 1],$$

$\alpha = 1$ corresponds to risk-neutral utility function.

The upper bound for $k = 1$ is $\overline{B}_1(v) = v$. To find the upper bound for $k = 2$, I need to compare the best case scenario given by winning the good w.p.1, and the worst case scenario given by beliefs that the opponents are bidding on the upper bound $\overline{B}_1(v)$. If a player with a private value v bids b and wins w.p.1, the expected payoff is $u(v - b) = (v - b)^{\alpha}$. In the worst case scenario a bidder solve the maximization problem:

$$\pi^*(v) = \max_{b \geq 0} u(v - b)F_0(b) = \max_{b \geq 0}(v - b)^{\alpha} F_0(b).$$

The best response $b^*(v)$ is a solution to the equation:

$$v - b^*(v) = \alpha \frac{F_0(b^*(v))}{f_0(b^*(v))}.$$

Thus, the upper bound for $k = 2$ is:

$$\overline{B}_2(v) = v - \alpha \frac{F_0^{\frac{1}{\alpha}+1}(b^*(v))}{f_0(b^*(v))}. \tag{6.5}$$

When α decreases to 0, bidders become more risk-averse, and their best-response to the worst case scenario $b^*(v)$ converges to the private value v while the upper bound $\overline{B}_2(v)$ converges to 45-degree line[29]. So, if the parameter of relative risk-aversion $r = 1 - \alpha$ is not bounded away

[29] It is a well-known fact that risk-averse players overbid in equilibrium with respect to risk-neutral equilibrium bidding function (see Krishna (2010)). The intuition is that risk-averse bidder buy "insurance" from losing the auction. The same logic is true in case of k–rationalizable bounds, i.e. bounds for a fixed k is larger if bidders are more risk-averse.

from 1, bounds for any fixed k can be arbitrary close to 45-degree line and the only fact that can be tested is that players bid lower than their private value, i.e.:

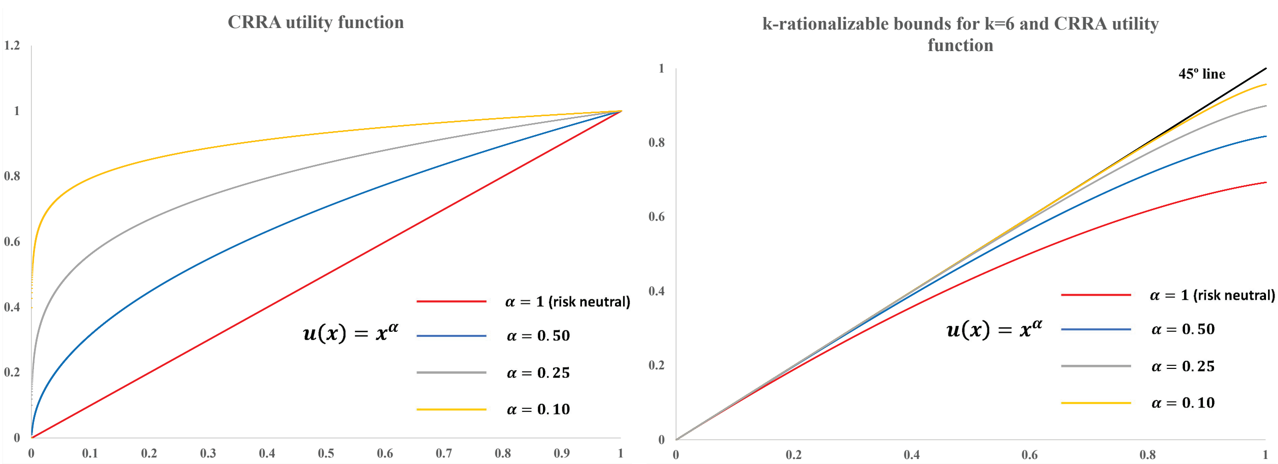
$$F_0(t) \leq G(t).$$

If no auxiliary data is available and no assumptions on the bidders risk-aversion is imposed then, k–rationalizability model can not be rejected for any given k. Consider an example with the distribution of private values $F_0(v) = v^{\theta}, \theta > 0$. Plugging it in (6.5):

$$\overline{B}_2(v) = v - \frac{\alpha \theta^{\frac{\theta}{\alpha}}}{(\theta + \alpha)^{\frac{\alpha+\theta}{\alpha}}} v^{\frac{\alpha+\theta}{\alpha}}.$$

When α decreases to zero and θ increases to infinity, the upper bound converges to 45-degree line. Similar to the result 4.1, it can be shown that for any distribution of bids $G(t)$ there exist $\alpha \in (0, 1]$ and $\theta > 0$ large enough such that bidders behavior can be k-rationalizable for any given k. This is shown in Figure 2.

Figure 2: $k-$rationalizable bounds with risk-aversion (for $k = 6$). An illustration with CRRA utility function and $U[0, 1]$ valuations.



6.5 Inference when there is collusion in ascending auctions

Recent results by Kaplan et al. (2017) show conditions under which the marginal distribution of bidders' values can be nonparametrically identified from ascending auctions data when there is collusion. Their maintained assumptions are the following:

1. Bidders draw their values independently (asymmetry is allowed).
2. Transaction price and all losing bids are observed.

3. Ascending auctions have a button-auction format, so losing bids (dropout prices) correspond exactly to bidders' valuations.

4. There is at least one known competitive bidder.

5. There is a set of potential colluders, but collusion is efficient, meaning that the cartel leader is always the one who draws the highest valuation.

Also implicit in their setup is the assumption of a nonbinding transaction price. Under these conditions, using de-censoring techniques Kaplan et al. (2017) show that the distribution of bidders' values can be nonparametrically identified. If I have access to the kind of rich data that is required, I can use their results to nonparametrically estimate the distribution of bidders' values in a way that is robust to the presence of collusion. However, the $k-$rationalizable bounds are constructed under the assumption of competitive bidding in first-price auctions.

6.6 Testing when the assumption $F_{0,1} = F_{0,2}$ is misspecified

Let $F_{0,1}(\cdot|x)$ and $F_{0,2}(\cdot|x)$ denote the distribution of values in first-price and ascending auctions, respectively. A key assumption in the construction of the nonparametric test was that $F_{0,1}(\cdot|x) = F_{0,2}(\cdot|x)$. However, as it was pointed out in Section 6.1, in order to construct an upper bound for $\overline{B}_k(v|x)$, *all that is needed is a valid lower bound for* $F_{0,1}(\cdot|x)$. Therefore, if the model is misspecified and $F_{0,1}(\cdot|x) \neq F_{0,2}(\cdot|x)$, a nonparametric test based on $\widehat{F}_{0,2}(\cdot|x)$ would still control for size (but may be conservative) if $F_{0,2}(\cdot|x) \leq F_{0,1}(\cdot|x)$ for a.e $x \in \mathscr{X}$. This would occur if more valuable objects are sold through ascending auctions.

Alternatively, I may relax the assumption that $F_{0,1}(\cdot|x) = F_{0,2}(\cdot|x)$ to the weaker condition that there exist $\mathcal{X}$ and $\mathcal{X}'$ such that

$$x \in \mathcal{X},\, x' \in \mathcal{X}' \implies F_{0,2}(\cdot|x) \leq F_{0,1}(\cdot|x'). \tag{6.6}$$

Let $v_{k,2}$ denote the inverse of the $k-$rationalizable upper bound that would obtain if I plug in $F_{0,2}$ in place of $F_{0,1}$ and let $v_{k,1}$ correspond to $F_{0,1}$. Suppose bids are $k-$rationalizable. The same arguments that led to (6.1) would yield now

$$F_{0,2}\left(v_{k,2}(t|x')|x'\right) \leq F_{0,1}\left(v_{k,1}(t|x)|x\right) \leq G_1(t|x),$$

where the first inequality follows from the condition described above and the second one follows from $k-$rationalizability. Define $\phi_k(t|x, x') = F_{0,2}\left(v_{k,2}(t|x')|x'\right) - G_1(t|x)$. The restriction of $k-$rationalizability can be based on the test

$$\phi_k(t|x, x') \cdot 1\left[x \in \mathcal{X},\, x' \in \mathcal{X}'\right] \leq 0, \quad \forall\ (x, x', t).$$

A condition like (6.6) can arise, for example, if I make additional assumptions about how the distribution of values shifts with particular elements in x. One instance in which this may happen involves endogenous participation.

6.6.1 Endogenous participation

Let $F_{0,1}(\cdot|X,n)$ and $F_{0,2}(\cdot|X,n)$ denote the distribution of values in first-price and ascending auctions respectively, where n is a number of participants. As was pointed out in Remark 3.1, the most natural model of bidder participation that can be reconciled with the assumption that $F_{0,1}(\cdot|X,n) = F_{0,2}(\cdot|X,n)$ is that participation is exogenous. If the assumption of the exogenous participation is violated such that more valuable objects attract more participants and more valuable objects are sold through ascending auctions, then we can relax the assumption $F_{0,1} = F_{0,2}$ to the weaker condition,

$$F_{0,2}(\cdot|x,n) \leq F_{0,1}(\cdot|x,n'), \quad \forall n > n'.$$

To test k–rationalizability in auctions with n' participants under this special case of endogenous entry I can plug in $F_{0,2}(\cdot|x,n)$ for any $n > n'$ instead of $F_0(\cdot|x,n')$ in (3.4). The test for k–rationalizability is still valid but can be conservative.

7 Concluding remarks

In this paper I analyze testable implications of strategically sophisticated bidding in first-price auctions without assumption of equilibrium behavior. The model of $k-$ rationalizability I use is consistent with many possible patterns of deviation from BNE: it allows for overbidding or underbidding with respect to risk-neutral BNE, as well as for heterogeneity in beliefs and for the possibility that beliefs depend on bidders' observed signals (values). Importantly, it includes BNE and (full) rationalizability as special cases, but it also allows for finitely many steps of deletion of strategies that the not best responces. $k-$rationalizable bidding functions are completely characterized by an upper bound, which leads to stochastic dominance implications. I propose tests that identify the largest value of k such that bidding behavior in the population is consistent with the properties of $k-$rationalizability. This can be done assuming symmetry or by dividing bidders into "types" according to observable characteristics. Rejecting a finite k implies that there are bidders who violate the $k-$rationalizable bounds with positive probability. It automatically rejects BNE as well as full rationalizability as the true underlying models. The test quantifies the extent of the deviation from rational bidding behavior. To the best of my knowledge, the methodology proposed in this paper constitutes the most robust tests of rationalizability in first-price auctions. Relying exclusively on bounds for $k-$rationalizable bids, I avoid any assumptions about the existence of "behavioral types", which characterize models of cognitive hierarchy or "Level-k thinking".

I show that under the IPV assumption and with access to auxiliary data from ascending auctions from the same population of bidders a stochastic dominance test for $k-$ rationalizability can be performed nonparametrically. I proposed a testing procedure that adapts to the properties of the contact sets. It leads to non-conservative results vis-a-vis methods based on "least-favorable configurations". The test has asymptotically pivotal features and is computationally easy to implement. I also proposed similar one-sided tests for BNE which allows to test risk-neutral BNE and risk-averse BNE. I discuss extensions to semi-parametric models which can be used, for example, when there is no auxiliary data from ascending auctions or when values are interdependent in first-price auctions. In other extensions I considered cases where the distribution of values is only partially identified, where bidders have ambiguity about such distribution or when there is presence of collusive behavior in the population of ascending auctions.

As an illustration, I apply my methodology to USFS timber auctions in the Pacific Northwest (Oregon and Washington) during the period 1994-2007. My data combines ascending and first-price auctions which facilitates the nonparametric implementation of my procedure. Dividing bidders into "small" and "large" types according to the number of employees and focusing on auctions with two participants, I find that values of k as low as $k = 2$ are rejected in asymmetric auctions, whereas no value of k can be rejected in symmetric auctions. However, BNE was rejected in all cases. In the case of asymmetric auctions, my results suggested that large bidders routinely overestimate the bids that are submitted by small bidders, which can lead to economic losses in expected profits. Using counterfactual analysis I find that the median improvement in large firms' expected profits from a simple correction towards the most pessimistic reasonable beliefs is about 4%, but this improvement was estimated to be greater than 50% in 25% of the auctions in the sample. The counterfactual analysis also hints at the presence of substantial overbidding by large bidders, particularly in auctions involving the most valuable tracts.

References

An, Y. (2017). Identification of first-price auctions with non-equilibrium beliefs: A measurement error approach. *Journal of Econometrics.*

Andrews, D. and Shi, X. (2011). Nonparametric inference based on conditional moment inequalities. Working Paper. Yale University.

Andrews, D. and Shi, X. (2013). Inference based on conditional moment inequality models. *Econometrica*, 81(2):609–666.

Aradillas-Lopez, A., Gandhi, A., and Quint, D. (2016). A simple test for moment inequality models with an application to english auctions. *Journal of Econometrics*, 194:96 – 115.

Aradillas-López, A., Gandhi, A., and Quint, D. (2016). A simple test for moment inequality models with an application to english auctions. *Journal of Econometrics*, 194(1):96–115.

Aradillas-Lopez, A. and Tamer, E. (2008). The identification power of equilibrium in simple games. *Journal of Business & Economic Statistics*, 26(3):261–283.

Aryal, G., Grundl, S., Kim, D.-H., and Zhu, Y. (2016). Empirical relevance of ambiguity in first-price auctions.

Athey, S. and Haile, P. A. (2002). Identification of standard auction models. *Econometrica*, 70(6):2107–2140.

Athey, S., Levin, J., and Seira, E. (2011). Comparing open and sealed bid auctions: Evidence from timber auctions. *The Quarterly Journal of Economics*, 126(1):207–257.

Baldwin, L., Marshall, R., and Richard, J. (1997). Bidder collusion at forest service timber auctions. *Journal of Political Economy*, 105:657–699.

Battigalli, P. and Siniscalchi, M. (2003). Rationalizable bidding in first-price auctions. *Games and Economic Behavior*, 45(1):38–72.

Bernheim, D. (1984). Rationalizable strategic behavior. *Econometrica*, 52:1007–1028.

Chernozhukov, V., Lee, S., and Rosen, A. (2013). Intersection bounds: Estimation and inference. *Econometrica*, 81(2):667–738.

Crawford, V. P., Costa-Gomes, M., , and Iriberri, N. (2013). Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications. *Journal of Economic Literature*, 51(1):5–62.

Crawford, V. P. and Iriberri, N. (2007). Level-k auctions: Can a nonequilibrium model of strategic thinking explain the winner's curse and overbidding in private-value auctions? *Econometrica*, 75(6):1721–1770.

David, H. and Nagaraja, H. (2003). *Order Statistics*. Wiley Series in Probability and Statistics.

Donald, S. and Paarsch, H. (1996). Identification, estimation, and testing in parametric empirical models of auctions within the independent private values paradigm. *Econometric Theory*, 12:517–567.

Dwas, M. (1956). The large-sample power of rank tests in the two-sample problem. *Annals of Mathematical Statistics*, 27:352–374.

Fudenberg, D. and Tirole, J. (1998). *Game Theory*. MIT Press. Cambridge, MA.

Gillen, B. (2009). Identification and estimation of level-k auctions.

Guerre, E., Perrigne, I., and Vuong, Q. (2000). Optimal nonparametric estimation of first-price auctions. *Econometrica*, 68(3):525–574.

Haile, P., Hong, H., and Shum, M. (2004). Nonparametric tests for common values in first-price sealed-bid auctions. Working Paper. Yale University.

Haile, P. A. and Tamer, E. (2003). Inference with an incomplete model of english auctions. *Journal of Political Economy*, 111(1):1–51.

Hendricks, K., Pinkse, J., and Porter, R. H. (2003). Empirical implications of equilibrium bidding in first-price, symmetric, common value auctions. *The Review of Economic Studies*, 70(1):115–145.

Hendricks, K. and Porter, R. (1988). An empirical study of an auction with asymmetric information. *American Economic Review*, 78:865–883.

Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Annals of Mathematical Statistics*, 19:293–325.

Hortacsu, A., Luco, F., Puller, S. L., and Zhu, D. (2017). Does strategic ability affect efficiency? evidence from electricity markets.

Hubbard, T., Li, T., and Paarsch, H. (2012). Semiparametric estimation in models of first-price, sealed-bid auctions with affiliation. *Journal of Econometrics*, 168:4–16.

Kagel, J. H. and Levin, D. (2014). Auctions: A survey of experimental research'.

Kaplan, U., Marmer, V., and Shneyerov, A. (2017). Identifying collusion in english auctions. Working Paper, University of British Columbia.

Kline, B. (2015). An empirical model of non-equilibrium behavior in games.

Kosorok, M. (2008). *Introduction to Empirical Processes and Semiparametric Inference*. Springer.

Krishna, V. (2010). *Auction Theory*. Academic Press.

Laffont, J.-J., Ossard, H., and Vuong, Q. (1995). Econometrics of first-price auctions. *Econometrica*, 63(4):953–980.

Lee, S., Song, K., and Whang, Y. (2013). Testing functional inequalities. *Journal of Econometrics*, 172(1):14 – 32.

Lee, S., Song, K., and Whang, Y. (2014). Testing for a general class of functional inequalities. Working Paper.

Lehmann, E. (1951). Consistency and unbiasedness of certain nonparametric tests. *Annals of Mathematical Statistics*, 22:165–179.

Li, T., Perrigne, I., and Vuong, Q. (2000). Conditionally independent private information in ocs wildcat auctions. *Journal of Econometrics*, 98:129–161.

Li, T., Perrigne, I., and Vuong, Q. (2002). Structural estimation of the affiliated private value auction model. *RAND Journal of Economics*, 33:171–193.

Lu, J. and Perrigne, I. (2008). Estimating risk aversion from ascending and sealed-bid auctions: the case of timber auction data. *Journal of Applied Econometrics*, 23(7):871–896.

Milgrom, P. R. and Weber, R. J. (1982). A theory of auctions and competitive bidding. *Econometrica: Journal of the Econometric Society*, pages 1089–1122.

Osborne, M. and Rubinstein, A. (1994). *A Course in Game Theory*. MIT Press. Cambridge, MA.

Paarsch, H. J. (1992). Deciding between the common and private value paradigms in empirical models of auctions. *Journal of econometrics*, 51(1):191–215.

Pakes, A. and Pollard, D. (1989). Simulation and the asymptotics of optimization estimators. *Econometrica*, 57(5):1027–1057.

Pearce, D. (1984). Rationalizable strategic behavior and the problem of perfection. *Econometrica*, 52:1029–1050.

Pinkse, J. and Tan, G. (2005). The affiliation effect in first-price auctions. *Econometrica*, 73(1):263–277.

Pollard, D. (1990). Empirical processes: theory and applications. In *NSF-CBMS regional conference series in probability and statistics*, pages i–86. JSTOR.

Porter, R. and Zona, D. (1993). Detection of bid rigging in procurement auctions. *Journal of Political Economy*, 101(3):518–538.

Schurter, K. (2017). Identification and inference in first-price auctions with collusion. Working Paper.

Serfling, R. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley Series in Probability and Statistics.

Sherman, R. (1994). Maximal inequalities for degenerate u-processes with applications to optimization estimators. *Annals of Statistics*, 22:439–459.

Somaini, P. (2015). Identification in auction models with interdependent costs. Working Paper. Stanford GSB.

A Econometric Appendix

A.1 Proof of Theorem 3.1. Preliminary results

Before proving Theorem 3.1, I will first characterize the relevant asymptotic properties of the nonparametric estimators $\widehat{G}_1(t|x)$ and $\widehat{F}_0(\widehat{\overline{v}}_k(t|x)|x)$. The testing range sets $\mathscr{X}$ and $\mathscr{B}$ are those described in Assumption 3.4.

A.1.1 An asymptotic linear representation result for $\widehat{\mathbf{G}}_1(\mathbf{t}|\mathbf{x}) - \mathbf{G}_1(\mathbf{t}|\mathbf{x})$

Take the marginal kernel k described in Assumption 3.9. Recall that the support of the kernel is $[-S, S]$. For a given $u \in \mathbb{R}$ we will denote $\delta(u) = \int_{-S}^{u} k(\psi)d\psi$. Note that $\delta(u) = 1$ if $u \geq S$ and $\delta(u) = 0$ if $u < -S$. Also note that $\underset{u\in\mathbb{R}}{sup}\,|\delta(u)| \leq \int_{-S}^{S} |k(\psi)|\, d\psi < \infty$. Now, for bid b_j^i in first-price auction j and a fixed b let

$$m(b_j^i, b; h_{b,L_1}) = 1\left[b_i^j \leq b - h_{b,L_1}\cdot S\right] + \delta\left(\frac{b - b_j^i}{h_{p,L_2}}\right)\cdot 1\left[b - h_{b,L_1}\cdot S < b_j^i < b + h_{b,L_1}\cdot S\right].$$

Note that

$$m(b_j^i, b; h_{b,L_1}) = \frac{1}{h_{b,L_1}}\int_{-\infty}^{b} k\left(\frac{b - b_j^i}{h_{b,L_1}}\right) db,$$

and under the conditions in Assumptions 3.6 and 3.9,

$$\sup_{\substack{b\in\mathscr{B}(x)\\ x\in\mathscr{X}}}\left|E\left[m(b_j^i,b;h_{b,L_1})|X_j=x\right]-G_1(b|x)\right|=O\left(h_{b,L_1}\right).$$

As in the main text, group all the bids submitted in the j^{th} first-price auction as $\boldsymbol{b}_j$, and in an auction with n bids submitted, define

$$\overline{m}_n\left(\boldsymbol{b}_j,b;h_{b,L_1}\right)=\frac{1}{n}\sum_{i=1}^{n}m(b_j^i,b;h_{b,L_1}).$$

Fix x and, as always, let n denote the element in x corresponding to number of participating bidders. Fix b, and define

$$\widehat{R}_1(b|x)=\frac{1}{L_1h_{x,L_1}^c}\sum_{j=1}^{L_1}\mathcal{H}(X_j-x;h_{x,L_1})\cdot\overline{m}_n\left(\boldsymbol{b}_j,b;h_{b,L_1}\right).$$

The estimator $\widehat{G}_1(b|x)$ described in Section 3.2.3 can be expressed as $\widehat{G}_1(b|x)=\widehat{R}_1(b|x)/\widehat{f}_{1,X}(x)$. Fix x and, once again, let n denote the element in x corresponding to number of participating bidders. Fix b, and define

$$\begin{aligned}
\varphi^{R_1}\left(\boldsymbol{b}_j,X_j,b,x;h_{b,L_1},h_{x,L_1}\right)=&\mathcal{H}(X_j-x;h_{x,L_1})\cdot\overline{m}_n(\boldsymbol{b}_j,b;h_{b,L_1})-E\left[\mathcal{H}(X_j-x;h_{x,L_1})\cdot\overline{m}_n(\boldsymbol{b}_j,b;h_{b,L_1})\right],\\
\varphi^{f_{1,x}}(X_j,x;h_{x,L_1})=&\mathcal{H}(X_j-x;h_{x,L_1})-E\left[\mathcal{H}(X_j-x;h_{x,L_1})\right],\\
\varphi^{G_1}\left(\boldsymbol{b}_j,X_j,b,x;h_{b,L_1},h_{x,L_1}\right)=&\frac{1}{f_{1,X}(x)}\cdot\varphi^{R_1}\left(\boldsymbol{b}_j,X_j,b,x;h_{b,L_1},h_{x,L_1}\right)-\frac{G_1(b|x)}{f_{1,x}(x)}\cdot\varphi^{f_{1,x}}(X_j,x;h_{x,L_1})
\end{aligned}$$

Note that $E\left[\varphi^{G_1}\left(\boldsymbol{b}_j,X_j,b,x;h_{b,L_1},h_{x,L_1}\right)\right]=0$. Let $\mathscr{B}$ be any compact subset of $int([\underline{b},\overline{b}])$. A second-order approximation along with the conditions in Assumptions 3.6 and 3.9 produce the following result,

$$\widehat{G}_1(b|x)-G_1(b|x)=\frac{1}{L_1h_{x,L_1}^c}\sum_{j=1}^{L_1}\varphi^{G_1}\left(\boldsymbol{b}_j,X_j,b,x;h_{b,L_1},h_{x,L_1}\right)+\xi_{L_1}^{G_1}(b,x),\tag{A.1}$$

$$\text{where}\quad\sup_{\substack{b\in\mathscr{B}\\ x\in\mathscr{X}}}\left|\xi_{L_1}^{G_1}(b,x)\right|=O\left(h_{b,L_1}\right)+O\left(h_{x,L_1}^M\right)=O\left(L_1^{-1/2-\epsilon}\right)\quad\text{for some }\epsilon>0.$$

The linear representation in (A.1) is the first key result for the analysis of my test. In the following section I will describe an analogous result for $\widehat{F}_0(\widehat{\overline{v}}_k(t|x)|x)-F_0(\overline{v}_k(t|x)|x))$.

A.1.2 An asymptotic linear representation result for $\widehat{F}_0(\widehat{\overline{v}}_k(t|x)|x) - F_0(\overline{v}_k(t|x)|x))$

Take the marginal kernel k described in Assumption 3.9. Recall that the support of the kernel is $[-S, S]$. For a given $u \in \mathbb{R}$ we will denote $\delta(u) = \int_{-S}^{u} k(\psi)d\psi$. Note that $\delta(u) = 1$ if $u \geq S$ and $\delta(u) = 0$ if $u < -S$. Now, for a given v let

$$m(P_i, v; h_{p,L_2}) = 1\left[P_i \leq v - h_{p,L_2} \cdot S\right] + \delta\left(\frac{v - P_i}{h_{p,L_2}}\right) \cdot 1\left[v - h_{p,L_2} \cdot S < P_i < v + h_{p,L_2} \cdot S\right].$$

Note that

$$m(P_i, v; h_{p,L_2}) = \frac{1}{h_{p,L_2}} \int_{-\infty}^{v} k\left(\frac{t - P_i}{h_{p,L_2}}\right) dt.$$

And since the kernel k is symmetric around zero, we can express

$$\widehat{F}_{2,P|X}(v|x) = \int_{-\infty}^{v} \widehat{f}_{2,P|X}(t|x)dt = \frac{1}{\widehat{f}_{2,X}(x)} \cdot \frac{1}{L_2 h_{x,L_2}^c} \sum_{i=1}^{L_2} m(P_i, v; h_{p,L_2}) \cdot \mathcal{H}\left(X_i - x; h_{x,L_2}\right).$$

For a given (v, x), define

$$\varphi^{R_{2,P|X}}\left(P_i, X_i, v, x; h_{p,L_2}, h_{x,L_2}\right) = m(P_i, v; h_{p,L_2}) \cdot \mathcal{H}(X_i - x; h_{x,L_2}) - E\left[m(P_i, v; h_{p,L_2}) \cdot \mathcal{H}(X_i - x; h_{x,L_2})\right],$$

$$\varphi^{f_{2,x}}\left(X_i, x; h_{x,L_2}\right) = \mathcal{H}(X_i - x; h_{x,L_2}) - E\left[\mathcal{H}(X_i - x; h_{x,L_2})\right],$$

$$\varphi^{F_{2,P|X}}\left(P_i, X_i, v, h; h_{p,L_2}, h_{x,L_2}\right) = \frac{1}{f_{2,x}(x)} \cdot \varphi^{R_2}\left(P_i, X_i, v, x; h_{p,L_2}, h_{x,L_2}\right) - \frac{F_{2,P|X}(v|x)}{f_{2,x}(x)} \cdot \varphi^{f_{2,x}}\left(X_i, x; h_{x,L_2}\right)$$

Note that $E\left[\varphi^{F_{2,P|X}}\left(P_i, X_i, v, h; h_{p,L_2}, h_{x,L_2}\right)\right] = 0$. Let $\mathscr{V}$ be any compact subset of $int([\underline{v}, \overline{v}])$. A second-order approximation along with the conditions in Assumptions 3.6 and 3.9 yield

$$\widehat{F}_{2,P|X}(v|x) - F_{2,P|X}(v|x) = \frac{1}{L_2 h_{x,L_2}^c} \sum_{i=1}^{L_2} \varphi^{F_{2,P|X}}\left(P_i, X_i, v, x; h_{p,L_2}, h_{x,L_2}\right) + \xi_{L_2}^{F_{2,P|X}}(v, x), \quad \text{(A.2)}$$

$$\textit{where} \quad \sup_{\substack{v \in \mathscr{V} \\ x \in \mathscr{X}}} \left|\xi_{L_2}^{F_{2,P|X}}(v, x)\right| = O\left(h_{p,L_2}\right) + O\left(h_{x,L_2}^M\right) = O\left(L_2^{-1/2-\epsilon}\right) \quad \text{for some } \epsilon > 0.$$

$$\sup_{\substack{v \in \mathscr{V} \\ x \in \mathscr{X}}} \left|\widehat{F}_{2,P|X}(v|x) - F_{2,P|X}(v|x)\right| = O_p\left(\frac{1}{\sqrt{L_2^{1-\delta} h_{x,L_2}^c}}\right) \quad \forall\ \delta > 0.$$

Therefore, from our bandwidth convergence restrictions we get

$$\sup_{\substack{v \in \mathscr{V} \\ x \in \mathscr{X}}} \left|\widehat{F}_{2,P|X}(v|x) - F_{2,P|X}(v|x)\right| = O_p\left(L_2^{-1/4-\epsilon/2}\right) \quad \text{for some } \epsilon > 0.$$

Next, fix x and let n denote the value that corresponds to number of participating bidders in x.

Let

$$\Gamma^{(1)}\left(F_{2,P|X}(v|x)\right)=\frac{1}{n(n-1)F_{2,P|X}(v|x)^{n-2}\left(1-F_{2,P|X}(v|x)\right)} \tag{A.3}$$

$$\varphi^{F_0}\left(P_i,X_i,v,x;h_{p,L_2},h_{x,L_2}\right)=\Gamma^{(1)}\left(F_{2,P|X}(v|x)\right)\cdot\varphi^{F_{2,P|X}}\left(P_i,X_i,v,x;h_{p,L_2},h_{x,L_2}\right).$$

Note that $E\left[\varphi^{F_0}\left(P_i,X_i,v,x;h_{p,L_2},h_{x,L_2}\right)\right]=0$. By the definition of $\widehat{F}_0(v|x)$ and the result in (A.2),

$$\begin{gathered}\widehat{F}_0(v|x)-F_0(v|x)=\frac{1}{L_2h^c_{x,L_2}}\sum_{i=1}^{L_2}\varphi^{F_0}\left(P_i,X_i,v,x;h_{p,L_2},h_{x,L_2}\right)+\xi^{F_0}_{L_2}(v,x),\\ \text{where}\quad \underset{\substack{v\in\mathscr{V}\\x\in\mathscr{X}}}{sup}\left|\xi^{F_0}_{L_2}(v,x)\right|=O_p\left(L_2^{-1/2-\epsilon}\right)\quad\text{for some }\epsilon>0.\end{gathered} \tag{A.4}$$

Asymptotic properties of $\widehat{F}_0\left(\widehat{\overline{v}}_k(t|x)|x\right)-F_0\left(\overline{v}_k(t|x)|x\right)$ for $k=2$

Fix x and let n the value that corresponds to number of participating bidders in x. Let

$$\pi_{k=2}(v,b,x)=(v-b)F_0^{n-1}(b|x).$$

Denote $\nabla_b\pi_{k=2}(v,b,x)\equiv\frac{\partial\pi_{k=2}(v,b,x)}{\partial b}$. We have,

$$\nabla_b\pi_{k=2}(v,b,x)=-F_0^{n-1}(b|x)+(n-1)\cdot(v-b)F_0^{n-2}(b|x)\Gamma^{(1)}\left(F_{2,P|X}(b|x)\right)f_{2,P|X}(b|x).$$

Let $b^*_{k=2}(v|x)=\underset{b\geq 0}{argmax}\ \pi_{k=2}(v,b,x)$. Under our assumptions, $b^*_{k=2}(v|x)$ is characterized by the first-order conditions

$$\nabla_b\pi_{k=2}\left(v,b^*_{k=2}(v|x),x\right)=0. \tag{A.5}$$

By definition of the $k-$rationalizable bounds, $\overline{B}_{k=2}(v|x)=v-\pi_{k=2}\left(v,b^*_{k=2}(v|x),x\right)$.
My estimator for $\overline{B}_{k=2}(v|x)$ is $\widehat{\overline{B}}_{k=2}(v|x)=v-\widehat{\pi}_{k=2}\left(v,\widehat{b}^*_{k=2}(v|x),x\right)$,
where $\widehat{b}^*_{k=2}(v|x)=\underset{b\geq 0}{argmax}\ \widehat{\pi}_{k=2}(v,b,x)$, with $\widehat{\pi}_{k=2}(v,b,x)=(v-b)\widehat{F}_0^{n-1}(b|x)$.
By the design of my estimator $\widehat{F}_0^{n-1}(b|x)$, the derivative $\nabla_b\widehat{\pi}_{k=2}(v,b,x)\equiv\frac{\partial\widehat{\pi}_{k=2}(v,b,x)}{\partial b}$ has the exact sample-analog structure of $\nabla_b\pi_{k=2}(v,b,x)$. Namely,

$$\nabla_b\widehat{\pi}_{k=2}(v,b,x)=-\widehat{F}_0^{n-1}(b|x)+(n-1)\cdot(v-b)\widehat{F}_0^{n-2}(b|x)\Gamma^{(1)}\left(\widehat{F}_{2,P|X}(b|x)\right)\widehat{f}_{2,P|X}(b|x),$$

and $\widehat{b}^*_{k=2}(v|x)$ satisfies the sample-analog first order conditions

$$\nabla_b\widehat{\pi}_{k=2}\left(v,\widehat{b}^*_{k=2}(v|x),x\right)=0. \tag{A.6}$$

The primary goal is to characterize the asymptotic properties of $\widehat{\overline{B}}_{k=2}(v|x)-\overline{B}_{k=2}(v|x)$. Let

$\mathscr{V}$ be any compact subset of $int([\underline{v},\overline{v}])$ and let $\mathscr{B}$ be any compact subset of $int([\underline{b},\overline{b}])$. Using our previous results and bandwidth convergence conditions in Assumption 3.9, we have $\sup_{\substack{v\in\mathscr{V},b\in\mathscr{B}\\x\in\mathscr{X}}} |\nabla_b\widehat{\pi}_{k=2}(v,b,x) - \nabla_b\pi_{k=2}(v,b,x)| = O_p\left(L_2^{-1/4-\epsilon/2}\right)$ for some $\epsilon > 0$. Combining this with (A.5) and (A.6), we also obtain $\sup_{\substack{v\in\mathscr{V}\\x\in\mathscr{X}}} \left|\widehat{b}^*_{k=2}(v|x) - b^*_{k=2}(v|x)\right| = O_p\left(L_2^{-1/4-\epsilon/2}\right)$ for some $\epsilon > 0$. From a second-order approximation we get

$$\widehat{\pi}_{k=2}\left(v, b^*_{k=2}(v|x), x\right) = \widehat{\pi}_{k=2}\left(v, \widehat{b}^*_{k=2}(v|x), x\right) \quad + \underbrace{\nabla_b\widehat{\pi}_{k=2}\left(v, \widehat{b}^*_{k=2}(v|x), x\right)}_{=0 \text{ from (A.6)}} \cdot \left(b^*_{k=2}(v|x) - \widehat{b}^*_{k=2}(v|x)\right)$$
$$+O_p\left(\left|\widehat{b}^*_{k=2}(v|x) - b^*_{k=2}(v|x)\right|^2\right).$$

Therefore,

$$\sup_{\substack{v\in\mathscr{V}\\x\in\mathscr{X}}} \left|\widehat{\pi}_{k=2}\left(v, \widehat{b}^*_{k=2}(v|x), x\right) - \widehat{\pi}_{k=2}\left(v, b^*_{k=2}(v|x), x\right)\right| = O_p\left(L_2^{-1/2-\epsilon}\right) \quad \text{for some } \epsilon > 0. \tag{A.7}$$

And from here,

$$\widehat{\overline{B}}_{k=2}(v|x) = \overline{B}_{k=2}(v|x) + \left(\widehat{\pi}_{k=2}\left(v, b^*_{k=2}(v|x), x\right) - \pi_{k=2}\left(v, b^*_{k=2}(v|x), x\right)\right) + \varrho_{L_2}^{\overline{B}_{k=2}}(v,x), \tag{A.8}$$

$$\text{where} \quad \sup_{\substack{v\in\mathscr{V}\\x\in\mathscr{X}}} \left|\varrho_{L_2}^{\overline{B}_{k=2}}(v,x)\right| = O_p\left(L_2^{-1/2-\epsilon}\right) \quad \text{for some } \epsilon > 0.$$

Let φ^{F_0} be as defined in (A.3). Now for a given v, x let

$$\begin{aligned}&\varphi^{\overline{B}_{k=2}}\left(P_i, X_i, v, x; h_{p,L_2}, h_{x,L_2}\right) = \\ &(n-1)\cdot F_0^{n-2}\left(b^*_{k=2}(v|x)|x\right)\cdot\left(v - b^*_{k=2}(v|x)\right)\cdot\varphi^{F_0}\left(P_i, X_i, b^*_{k=2}(v|x), x; h_{p,L_2}, h_{x,L_2}\right)\end{aligned} \tag{A.9}$$

Note that $E\left[\varphi^{\overline{B}_{k=2}}\left(P_i, X_i, v, x; h_{p,L_2}, h_{x,L_2}\right)\right] = 0$. Combining (A.4) and (A.8) obtains,

$$\widehat{\overline{B}}_{k=2}(v|x) - \overline{B}_{k=2}(v|x) = \frac{1}{L_2 h^c_{x,L_2}} \sum_{i=1}^{L_2} \varphi^{\overline{B}_{k=2}}\left(P_i, X_i, v, x; h_{p,L_2}, h_{x,L_2}\right) + \xi_{L_2}^{\overline{B}_{k=2}}(v,x), \tag{A.10}$$

$$\text{where} \quad \sup_{\substack{v\in\mathscr{V}\\x\in\mathscr{X}}} \left|\xi_{L_2}^{\overline{B}_{k=2}}(v,x)\right| = O_p\left(L_2^{-1/2-\epsilon}\right) \quad \text{for some } \epsilon > 0.$$

Fix x. As in the main text, let $\overline{v}_k(\cdot|x)$ and $\widehat{\overline{v}}_k(\cdot|x)$ denote the inverse functions of $\overline{B}_k(\cdot|x)$ and $\widehat{\overline{B}}_k(\cdot|x)$, respectively. That is, for a given t, $\overline{v}_{k=2}(\cdot|x)$ and $\widehat{\overline{v}}_{k=2}(\cdot|x)$ are given, respectively, by the solution (in v) to the equations

$$\underbrace{v-\left(v-b^*_{k=2}(v|x)\right)F_0^{n-1}\left(b^*_{k=2}(v|x)|x\right)}_{\overline{B}_{k=2}(v|x)}=t,\quad \underbrace{v-\left(v-\widehat{b}^*_{k=2}(v|x)\right)\widehat{F}_0^{n-1}\left(\widehat{b}^*_{k=2}(v|x)|x\right)}_{\widehat{\overline{B}}_{k=2}(v|x)}=t.$$

By the first-order conditions (A.5) and (A.6), the Envelope Theorem yields

$$\nabla_v\overline{B}_{k=2}(v|x)=1-F_0^{n-1}\left(b^*_{k=2}(v|x)|x\right),\quad \nabla_v\widehat{\overline{B}}_{k=2}(v|x)\ =1-\widehat{F}_0^{n-1}\left(\widehat{b}^*_{k=2}(v|x)|x\right).$$

By construction, $\widehat{\overline{B}}_{k=2}\left(\widehat{\overline{v}}_k(t|x)|x\right)=\overline{B}_{k=2}\left(\overline{v}_k(t|x)|x\right)$ (both expressions are equal to t). A second-order approximation on the left-hand side combined with the expressions for $\nabla_v\overline{B}_{k=2}(v|x)$ and $\nabla_v\widehat{\overline{B}}_{k=2}(v|x)$ and (A.10) yield the following. Fix t and x. Let $\varphi^{\overline{B}_{k=2}}$ be as described in (A.9) and define

$$\varphi^{\overline{v}_{k=2}}\left(P_i,X_i,t,x;h_{p,L_2},h_{x,L_2}\right)=-\frac{1}{1-F_0^{n-1}\left(b^*_{k=2}(\overline{v}_{k=2}(t|x)|x)|x\right)}\cdot\varphi^{\overline{B}_{k=2}}\left(P_i,X_i,\overline{v}_{k=2}(t|x),x;h_{p,L_2},h_{x,L_2}\right) \tag{A.11}$$

Note that $E\left[\varphi^{\overline{B}_{k=2}}\left(P_i,X_i,t,x;h_{p,L_2},h_{x,L_2}\right)\right]=0$. Our previous results yield,

$$\begin{aligned}&\widehat{\overline{v}}_{k=2}(t|x)-\overline{v}_{k=2}(t|x)=\frac{1}{L_2h^c_{x,L_2}}\sum_{i=1}^{L_2}\varphi^{\overline{v}_{k=2}}\left(P_i,X_i,t,x;h_{p,L_2},h_{x,L_2}\right)+\xi_{L_2}^{\overline{v}_{k=2}}(t,x),\\&\qquad\text{where}\quad \sup_{\substack{t\in\mathscr{B}\\x\in\mathscr{X}}}\left|\xi_{L_2}^{\overline{v}_{k=2}}(t,x)\right|=O_p\left(L_2^{-1/2-\epsilon}\right)\quad\text{for some }\epsilon>0.\end{aligned} \tag{A.12}$$

Fix t and x. Let φ^{F_0} and $\varphi^{\overline{v}_{k=2}}$ be as described in (A.3) and (A.11) and define

$$\begin{aligned}&\varphi^{F_{0,k=2}}\left(P_i,X_i,t,x;h_{p,L_2},h_{x,L_2}\right)=\varphi^{F_0}\left(P_i,X_i,\overline{v}_{k=2}(t|x),x;h_{p,L_2},h_{x,L_2}\right)\\&+\Gamma^{(1)}\left(F_{2,P|X}(\overline{v}_{k=2}(t|x)|x)\right)\cdot f_{2,P|X}(\overline{v}_{k=2}(t|x)|x)\cdot\varphi^{\overline{v}_{k=2}}\left(P_i,X_i,t,x;h_{p,L_2},h_{x,L_2}\right).\end{aligned} \tag{A.13}$$

Note once again that $E\left[\varphi^{F_{0,k=2}}\left(P_i,X_i,t,x;h_{p,L_2},h_{x,L_2}\right)\right]=0$. Using the previous results and a second-order approximation, we obtain

$$\begin{aligned}&\widehat{F}_0\left(\widehat{\overline{v}}_{k=2}(t|x)|x\right)-F_0\left(\overline{v}_{k=2}(t|x)|x\right)=\frac{1}{L_2h^c_{x,L_2}}\sum_{i=1}^{L_2}\varphi^{F_{0,k=2}}\left(P_i,X_i,t,x;h_{p,L_2},h_{x,L_2}\right)+\xi_{L_2}^{F_{0,k=2}}(t,x),\\&\qquad\text{where}\quad \sup_{\substack{t\in\mathscr{B}\\x\in\mathscr{X}}}\left|\xi_{L_2}^{F_{0,k=2}}(t,x)\right|=O_p\left(L_2^{-1/2-\epsilon}\right)\quad\text{for some }\epsilon>0.\end{aligned} \tag{A.14}$$

For the purposes of my econometric test, the linear representation in (A.14) is the most important result. Next I will describe inductively how it extends to $k\geq 3$.

Asymptotic properties of $\widehat{F}_0(\widehat{\overline{v}}_k(t|x)|x) - F_0\left(\overline{v}_k(t|x)|x\right)$ for $k \geq 3$

The steps will be analogous to the case $k = 2$. We now have

$$\pi_k(v,b,x) = (v-b)F_0^{n-1}\left(\overline{v}_{k-1}(b|x)|x\right), \quad \widehat{\pi}_k(v,b,x) \ = (v-b)\widehat{F}_0^{n-1}\left(\widehat{\overline{v}}_{k-1}(b|x)|x\right),$$

with $\quad b_k^*(v|x) = \underset{b\geq 0}{argmax}\ \pi_k(v,b,x), \quad \widehat{b}_k^*(v|x) = \underset{b\geq 0}{argmax}\ \widehat{\pi}_k(v,b,x).$

Suppose the estimator $\widehat{\overline{v}}_{k-1}$ is such that $\underset{\substack{b\in\mathscr{B}\\x\in\mathscr{X}}}{sup}\left|\widehat{\overline{v}}_{k-1}(b|x) - \overline{v}_{k-1}(b|x)\right|^2 = O_p\left(L_2^{-1/2-\epsilon}\right)$ for some $\epsilon > 0$, note that this is true for $k = 3$ by the result in (A.12). As we defined in Assumption 3.4, let

$$\mathscr{V}_k = \{v\colon\ v = \overline{v}_k(b|x) \text{ for some } b \in \mathscr{B} \text{ and } x \in \mathscr{X}.\}.$$

Then, by the conditions described in Assumption 3.4, a second-order approximation coupled with the first-order conditions of b_k^* and $\overline{b}_k^*$ and the Envelope Theorem yield a generalization of the result in (A.7). Namely,

$$\underset{\substack{v\in\mathscr{V}_k\\x\in\mathscr{X}}}{sup}\left|\widehat{\pi}_k\left(v,\widehat{b}_k^*(v|x),x\right) - \widehat{\pi}_k\left(v,b_k^*(v|x),x\right)\right| = O_p\left(L_2^{-1/2-\epsilon}\right) \quad \text{for some } \epsilon > 0. \tag{A.15}$$

This, in turn, leads to a generalization of (A.8),

$$\begin{aligned}\widehat{\overline{B}}_k(v|x) &= \overline{B}_k(v|x) + \Big(\widehat{\pi}_k\left(v,b_k^*(v|x),x\right) - \pi_k\left(v,b_k^*(v|x),x\right)\Big) + \varrho_{L_2}^{\overline{B}_k}(v,x),\\ \text{where} \quad &\underset{\substack{v\in\mathscr{V}_k\\x\in\mathscr{X}}}{sup}\left|\varrho_{L_2}^{\overline{B}_k}(v,x)\right| = O_p\left(L_2^{-1/2-\epsilon}\right) \quad \text{for some } \epsilon > 0.\end{aligned} \tag{A.16}$$

Next, suppose the estimator $\widehat{\overline{v}}_{k-1}$ satisfies a linear representation result of the form,

$$\begin{aligned}\widehat{\overline{v}}_{k-1}(t|x) - \overline{v}_{k-1}(t|x) &= \frac{1}{L_2h_{x,L_2}^c}\sum_{i=1}^{L_2}\varphi^{\overline{v}_{k-1}}\left(P_i,X_i,t,x;h_{p,L_2},h_{x,L_2}\right) + \xi_{L_2}^{\overline{v}_{k-1}}(t,x),\\ \text{where} \quad &\underset{\substack{t\in\mathscr{B}\\x\in\mathscr{X}}}{sup}\left|\xi_{L_2}^{\overline{v}_{k-1}}(t,x)\right| = O_p\left(L_2^{-1/2-\epsilon}\right) \quad \text{for some } \epsilon > 0.\end{aligned} \tag{A.17}$$

With $E\left[\varphi^{\overline{v}_{k-1}}\left(P_i,X_i,t,x;h_{p,L_2},h_{x,L_2}\right)\right] = 0$. Note that this has been established for $k = 3$ in (A.12). Let us go back to (A.16) and focus on the term $\widehat{\pi}_k\left(v,b_k^*(v|x),x\right) - \pi_k\left(v,b_k^*(v|x),x\right)$. This is given by

$$\widehat{\pi}_k\left(v, b_k^*(v|x), x\right) - \pi_k\left(v, b_k^*(v|x), x\right) = \left(v - b_k^*(v|x)\right) \times \left[\widehat{F}_0^{n-1}\left(\widehat{\overline{v}}_{k-1}(b_k^*(v|x)|x)|x\right) - F_0^{n-1}\left(\overline{v}_{k-1}(b_k^*(v|x)|x)|x\right)\right]$$

$$\begin{aligned} &= \left(v - b_k^*(v|x)\right) \times \Bigg\{\left[\widehat{F}_0^{n-1}\left(\widehat{\overline{v}}_{k-1}(b_k^*(v|x)|x)|x\right) - \widehat{F}_0^{n-1}\left(\overline{v}_{k-1}(b_k^*(v|x)|x)|x\right)\right] \\ &\qquad + \left[\widehat{F}_0^{n-1}\left(\overline{v}_{k-1}(b_k^*(v|x)|x)|x\right) - F_0^{n-1}\left(\overline{v}_{k-1}(b_k^*(v|x)|x)|x\right)\right]\Bigg\} \end{aligned} \tag{A.18}$$

Notice that the term $\left[\widehat{F}_0^{n-1}\left(\widehat{\overline{v}}_{k-1}(b_k^*(v|x)|x)|x\right) - \widehat{F}_0^{n-1}\left(\overline{v}_{k-1}(b_k^*(v|x)|x)|x\right)\right]$ was equal to zero in the case $k = 2$ because $\widehat{\overline{v}}_1(t|x) = \overline{v}_{k-1}(t|x) = t$, since $\widehat{\overline{B}}_{k=1}(v|x) = \overline{B}_{k=1}(v|x) = v$ (the upper bound for $k = 1$ is simply the 45-degree line and so is its inverse, neither of which has to be estimated). Let φ^{F_0} and $\varphi^{\overline{v}_{k-1}}$ be as described in (A.3) and (A.17), respectively. Fix v and x. As always, let n denote the value that corresponds to number of participating bidders in x. Define

$$\varphi^{\overline{B}_k}\left(P_i, X_i, v, x; h_{p,L_2}, h_{x,L_2}\right) = \tag{A.19}$$

$$(n-1) \cdot F_0^{n-2}\left(\overline{v}_{k-1}(b_k^*(v|x)|x)|x\right) \cdot \left(v - b_k^*(v|x)\right) \times \Bigg\{\varphi^{F_0}\left(P_i, X_i, \overline{v}_{k-1}(b_k^*(v|x)|x), x; h_{p,L_2}, h_{x,L_2}\right)$$

$$+ \Gamma^{(1)}\left(F_{2,P|X}\left(\overline{v}_{k-1}(b_k^*(v|x)|x)|x\right)\right) \cdot f_{2,P|X}\left(\overline{v}_{k-1}(b_k^*(v|x)|x)|x\right) \cdot \varphi^{\overline{v}_{k-1}}\left(P_i, X_i, b_k^*(v|x), x; h_{p,L_2}, h_{x,L_2}\right)\Bigg\}$$

Note that $E\left[\varphi^{\overline{B}_k}\left(P_i, X_i, v, x; h_{p,L_2}, h_{x,L_2}\right)\right] = 0$, as in all previous cases. The expression in (A.19) is a generalization of (A.9). Note that the second term on the right-hand side of (A.19) is absent in (A.9) because, as we pointed out above, $\widehat{\overline{v}}_1(t|x) = \overline{v}_1(t|x) = t$ (since the $k = 1$ bound is just the 45-degree line).

A second-order approximation to the first term in (A.18), combined with the results in (A.4) and (A.17) yield the following generalization of (A.10),

$$\begin{aligned} &\widehat{\overline{B}}_k(v|x) - \overline{B}_k(v|x) = \frac{1}{L_2 h_{x,L_2}^c} \sum_{i=1}^{L_2} \varphi^{\overline{B}_k}\left(P_i, X_i, v, x; h_{p,L_2}, h_{x,L_2}\right) + \xi_{L_2}^{\overline{B}_k}(v, x), \\ &\qquad \text{where} \quad \sup_{\substack{v \in \mathscr{V}_k \\ x \in \mathscr{X}}} \left|\xi_{L_2}^{\overline{B}_k}(v, x)\right| = O_p\left(L_2^{-1/2-\epsilon}\right) \quad \text{for some } \epsilon > 0. \end{aligned} \tag{A.20}$$

By the Envelope Theorem and the first-order conditions satisfied by $b_k^*(v|x)$ and $\widehat{b}_k^*(v|x)$, from the definition of $\overline{B}_k(v|x)$ and $\widehat{\overline{B}}_k(v|x)$ we obtain $\nabla_v \overline{B}_k(v|x) = 1 - F_0^{n-1}\left(b_k^*(v|x)|x\right)$ and $\nabla_v \widehat{\overline{B}}_k(v|x) = 1 - \widehat{F}_0^{n-1}\left(\widehat{b}_k^*(v|x)|x\right)$. Once again, let $\overline{v}_k(\cdot|x)$ and $\widehat{\overline{v}}_k(\cdot|x)$ denote the inverse functions of $\overline{B}_k(\cdot|x)$ and $\widehat{\overline{B}}_k(\cdot|x)$, respectively. By construction, for any t we must have $\widehat{\overline{B}}_k\left(\widehat{\overline{v}}_k(t|x)|x\right) =$

$\overline{B}_k\left(\overline{v}_k(t|x)|x\right)$ (since both expressions are equal to t, by definition). From here, we get

$$\widehat{\widehat{B}}_k\left(\widehat{\overline{v}}_k(t|x)|x\right)-\widehat{\widehat{B}}_k\left(\overline{v}_k(t|x)|x\right)=\overline{B}_k\left(\widehat{\overline{v}}_k(t|x)|x\right)-\widehat{\widehat{B}}_k\left(\overline{v}_k(t|x)|x\right).$$

The right-hand side of the above expression can be analyzed using (A.20). Combining this with a first-order approximation to the left hand side we obtain the following results, which are generalizations of (A.11) and (A.12). Let

$$\varphi^{\overline{v}_k}\left(P_i,X_i,t,x;h_{p,L_2},h_{x,L_2}\right)=-\frac{1}{1-F_0^{n-1}\left(b_k^*(\overline{v}_k(t|x)|x)|x\right)}\cdot\varphi^{\overline{B}_k}\left(P_i,X_i,\overline{v}_k(t|x),x;h_{p,L_2},h_{x,L_2}\right) \tag{A.21}$$

Note once again that $E\left[\varphi^{\overline{B}_k}\left(P_i,X_i,t,x;h_{p,L_2},h_{x,L_2}\right)\right]=0$. We have

$$\begin{gathered}\widehat{\overline{v}}_k(t|x)-\overline{v}_k(t|x)=\frac{1}{L_2h_{x,L_2}^c}\sum_{i=1}^{L_2}\varphi^{\overline{v}_k}\left(P_i,X_i,t,x;h_{p,L_2},h_{x,L_2}\right)+\xi_{L_2}^{\overline{v}_k}(t,x),\\ \text{where}\quad \underset{\substack{t\in\mathscr{B}\\x\in\mathscr{X}}}{sup}\left|\xi_{L_2}^{\overline{v}_k}(t,x)\right|=O_p\left(L_2^{-1/2-\epsilon}\right)\quad\text{for some }\epsilon>0.\end{gathered} \tag{A.22}$$

From here we can study the properties of $\widehat{F}_0\left(\widehat{\overline{v}}_k(t|x)|x\right)-F_0\left(\overline{v}_k(t|x)|x\right)$ straightforwardly. Fix t and x. Let φ^{F_0} and $\varphi^{\overline{v}_k}$ be as described in (A.3) and (A.21) and define

$$\begin{aligned}&\varphi^{F_{0,k}}\left(P_i,X_i,t,x;h_{p,L_2},h_{x,L_2}\right)=\varphi^{F_0}\left(P_i,X_i,\overline{v}_k(t|x),x;h_{p,L_2},h_{x,L_2}\right)\\&+\Gamma^{(1)}\left(F_{2,P|X}(\overline{v}_k(t|x)|x)\right)\cdot f_{2,P|X}(\overline{v}_k(t|x)|x)\cdot\varphi^{\overline{v}_k}\left(P_i,X_i,t,x;h_{p,L_2},h_{x,L_2}\right).\end{aligned} \tag{A.23}$$

Note once again that $E\left[\varphi^{F_{0,k}}\left(P_i,X_i,t,x;h_{p,L_2},h_{x,L_2}\right)\right]=0$ as in all previous cases. Using the previous results and a second-order approximation, we obtain

$$\begin{gathered}\widehat{F}_0\left(\widehat{\overline{v}}_k(t|x)|x\right)-F_0\left(\overline{v}_k(t|x)|x\right)=\frac{1}{L_2h_{x,L_2}^c}\sum_{i=1}^{L_2}\varphi^{F_{0,k}}\left(P_i,X_i,t,x;h_{p,L_2},h_{x,L_2}\right)+\xi_{L_2}^{F_{0,k}}(t,x),\\ \text{where}\quad \underset{\substack{t\in\mathscr{B}\\x\in\mathscr{X}}}{sup}\left|\xi_{L_2}^{F_{0,k}}(t,x)\right|=O_p\left(L_2^{-1/2-\epsilon}\right)\quad\text{for some }\epsilon>0.\end{gathered} \tag{A.24}$$

The linear representation in (A.24) extends the result in (A.14) to the case $k\geq 3$ and is key for the analysis of my test.

A.1.3 An asymptotic linear representation result for $\widehat{\phi}_k(t|x)-\phi_k(t|x)$

Combining our previous characterizations we obtain the result that will allow us to prove Theorem 3.1. Recall that

$$\widehat{\phi}_k(t|x) = \widehat{F}_0(\widehat{v}_k(t|x)|x) - \widehat{G}_1(t|x).$$

Also recall that I denoted $L \equiv L_1 + L_2$ as the combined sample sizes for first-price and ascending auctions. Let φ^{G_1} and $\varphi^{F_{0,k}}$ be as described in (A.1) and (A.24), respectively. From the results obtained there, we get

$$\begin{aligned}
\widehat{\phi}_k(t|x) - \phi_k(t|x) &= \frac{1}{L_2 h^c_{x,L_2}} \sum_{i=1}^{L_2} \varphi^{F_{0,k}}\left(P_i, X_i, t, x; h_{p,L_2}, h_{x,L_2}\right) \\
&- \frac{1}{L_1 h^c_{x,L_1}} \sum_{j=1}^{L_1} \varphi^{G_1}\left(\boldsymbol{b}_j, X_j, t, x; h_{b,L_1}, h_{x,L_1}\right) + \xi^{F_{0,k}}_{L_2}(t,x) - \xi^{G_1}_{L_1}(t,x) \\
&\text{where} \quad \begin{cases} \sup\limits_{\substack{t\in\mathscr{B}\\ x\in\mathscr{X}}} \left|\xi^{F_{0,k}}_{L_2}(t,x)\right| = O_p\left(L_2^{-1/2-\epsilon}\right), \\ \sup\limits_{\substack{t\in\mathscr{B}\\ x\in\mathscr{X}}} \left|\xi^{G_1}_{L_1}(t,x)\right| = O_p\left(L_1^{-1/2-\epsilon}\right). \end{cases} \quad \text{for some } \epsilon > 0.
\end{aligned} \tag{A.25}$$

A.2 Proof of Theorem 3.1. Final steps

Recall that

$$\widehat{\mathcal{T}}_k = \frac{1}{L_2} \sum_{i=1}^{L_2} \widehat{\Lambda}_{k,2}(X_i) W_{\mathscr{X}}(X_i) + \frac{1}{L_1} \sum_{i=1}^{L_1} \widehat{\Lambda}_{k,1}(X_i) W_{\mathscr{X}}(X_i).$$

Let us focus on the second term (the first term will have analogous properties using the same steps we will take next). We can decompose

$$\begin{aligned}
&\frac{1}{L_1} \sum_{i=1}^{L_1} \widehat{\Lambda}_{k,1}(X_i) W_{\mathscr{X}}(X_i) = \frac{1}{L_1} \sum_{i=1}^{L_1} \int_{t\in\mathscr{B}(X_i)} \widehat{\phi}_k(t|X_i) \cdot 1\left[\widehat{\phi}_k(t|X_i) \geq -\gamma_{L_1}\right] dQ(t|X_i) \cdot W_{\mathscr{X}}(X_i) \\
&= \frac{1}{L_1} \sum_{i=1}^{L_1} \int_{t\in\mathscr{B}(X_i)} \widehat{\phi}_k(t|X_i) \cdot 1\left[\phi_k(t|X_i) \geq 0\right] dQ(t|X_i) W_{\mathscr{X}}(X_i) \\
&+ \frac{1}{L_1} \sum_{i=1}^{L_1} \int_{t\in\mathscr{B}(X_i)} \phi_k(t|X_i) \cdot 1\left[\widehat{\phi}_k(t|X_i) \geq -\gamma_{L_1}\right] \cdot 1\left[\phi_k(t|X_i) < -2\gamma_{L_1}\right] dQ(t|X_i) \cdot W_{\mathscr{X}}(X_i) \\
&+ \frac{1}{L_1} \sum_{i=1}^{L_1} \int_{t\in\mathscr{B}(X_i)} \phi_k(t|X_i) \cdot 1\left[\widehat{\phi}_k(t|X_i) \geq -\gamma_{L_1}\right] \cdot 1\left[-2\gamma_{L_1} \leq \phi_k(t|X_i) < 0\right] dQ(t|X_i) \cdot W_{\mathscr{X}}(X_i) \\
&- \frac{1}{L_1} \sum_{i=1}^{L_1} \int_{t\in\mathscr{B}(X_i)} \phi_k(t|X_i) \cdot 1\left[\widehat{\phi}_k(t|X_i) < -\gamma_{L_1}\right] \cdot 1\left[\phi_k(t|X_i) \geq 0\right] dQ(t|X_i) \cdot W_{\mathscr{X}}(X_i) \\
&+ \frac{1}{L_1} \sum_{i=1}^{L_1} \int_{t\in\mathscr{B}(X_i)} \left(\widehat{\phi}_k(t|X_i) - \phi_k(t|X_i)\right) \cdot 1\left[\widehat{\phi}_k(t|X_i) \geq -\gamma_{L_1}\right] \cdot 1\left[\phi_k(t|X_i) < -2\gamma_{L_1}\right\} dQ(t|X_i) \cdot W_{\mathscr{X}}(X_i)
\end{aligned} \tag{A.26}$$

$$+\frac{1}{L_1}\sum_{i=1}^{L_1}\int_{t\in\mathscr{B}(X_i)}\left(\widehat{\phi}_k(t|X_i)-\phi_k(t|X_i)\right)\cdot 1\left[\widehat{\phi}_k(t|X_i)\geq-\gamma_{L_1}\right]\cdot 1\left[-2\gamma_{L_1}\leq\phi_k(t|X_i)<0\right]dQ(t|X_i)\cdot W_{\mathscr{X}}(X_i)$$

$$-\frac{1}{L_1}\sum_{i=1}^{L_1}\int_{t\in\mathscr{B}(X_i)}\left(\widehat{\phi}_k(t|X_i)-\phi_k(t|X_i)\right)\cdot 1\left[\widehat{\phi}_k(t|X_i)<-\gamma_{L_1}\right]\cdot 1\left[\phi_k(t|X_i)\geq 0\right]dQ(t|X_i)\cdot W_{\mathscr{X}}(X_i)$$

$$\equiv\frac{1}{L_1}\sum_{i=1}^{L_1}\int_{t\in\mathscr{B}(X_i)}\widehat{\phi}_k(t|X_i)\cdot 1\left[\phi_k(t|X_i)\geq 0\right]dQ(t|X_i)\cdot W_{\mathscr{X}}(X_i)+\frac{1}{L_1}\sum_{i=1}^{L_1}\int_{t\in\mathscr{B}(X_i)}\varepsilon_{k,1}(t|X_i)dQ(t|X_i),$$

Recall that $\int_{t\in\mathscr{B}(x)}dQ(t|x)=1$ for all x. Also note that $|\phi_k(t|x)|\leq 1$ for all t,x. Therefore,

$$\begin{aligned}
&\frac{1}{L_1}\sum_{i=1}^{L_1}\int_{t\in\mathscr{B}(X_i)}\varepsilon_{k,1}(t|X_i)dQ(t|X_i)\leq\frac{1}{L_1}\sum_{i=1}^{L_1}\int_{t\in\mathscr{B}(X_i)}\left|\varepsilon_{k,1}(t|X_i)\right|dQ(t|X_i)\\
&\leq 2\times 1\left[\sup_{\substack{t\in\mathscr{B}\\x\in\mathscr{X}}}\left|\widehat{\phi}_k(t|x)-\phi_k(t|x)\right|\geq\gamma_{L_1}\right]+2\gamma_{L_1}\times\sup_{t\in\mathscr{B}}\frac{1}{L_1}\sum_{i=1}^{L_1}1\left[-2\gamma_{L_1}\leq\phi_k(t|X_i)<0\right]\cdot W_{\mathscr{X}}(X_i)\\
&\qquad+2\cdot\sup_{\substack{t\in\mathscr{B}\\x\in\mathscr{X}}}\left|\widehat{\phi}_k(t|x)-\phi_k(t|x)\right|\cdot 1\left[\sup_{\substack{t\in\mathscr{B}\\x\in\mathscr{X}}}\left|\widehat{\phi}_k(t|x)-\phi_k(t|x)\right|\geq\gamma_{L_1}\right]\\
&\qquad+\sup_{\substack{t\in\mathscr{B}\\x\in\mathscr{X}}}\left|\widehat{\phi}_k(t|x)-\phi_k(t|x)\right|\cdot\sup_{t\in\mathscr{B}}\frac{1}{L_1}\sum_{i=1}^{L_1}1\left[-2\gamma_{L_1}\leq\phi_k(t|X_i)<0\right]W_{\mathscr{X}}(X_i)
\end{aligned}\tag{A.27}$$

I will analyze each term on the right-hand side of (A.27). First, notice that the regularity and smoothness conditions in Assumptions 3.4 and 3.6, combined with the bounded-variation properties of the kernel described in Assumption 3.9 imply, via Lemmas 2.4, 2.12, 2.13, 2.14 and Example 2.10 in Pakes and Pollard (1989), that the following classes of functions are Euclidean (see Definition 2.7 in Pakes and Pollard (1989)) for a constant envelope,

$$\begin{aligned}
\mathscr{F}_k&=\left\{\varphi^{F_{0,k}}(\cdot,\cdot,t,x;h_1,h_2)\colon t\in\mathscr{B},\,x\in\mathscr{X},\,h_1>0,\,h_2>0\right\},\\
\mathscr{G}&=\left\{\varphi^{G_1}(\cdot,\cdot,t,x;h_1,h_2)\colon t\in\mathscr{B},\,x\in\mathscr{X},\,h_1>0,\,h_2>0\right\}
\end{aligned}$$

Define the following two empirical processes $\nu_{L_1}(\cdot)$ and $\nu_{L_2}(\cdot)$ indexed over $\mathscr{B}\times\mathscr{X}$ as

$$\nu_{L_1}(t,x)=\frac{1}{L_1}\sum_{i=1}\varphi^{G_1}\left(\boldsymbol{b}_i,X_i,t,x;h_{b,L_1},h_{x,L_1}\right),\quad\nu^k_{L_2}(t,x)\ =\frac{1}{L_2}\sum_{i=1}\varphi^{F_{0,k}}\left(P_i,X_i,t,x;h_{b,L_1},h_{x,L_1}\right).$$

The Euclidean property of the above classes of functions, combined with Corollary 4 and the Main Corollary in Sherman (1994) imply that there exists a constant $\overline{D}$ such that, for any $\delta>0$,

$$Pr\left[\sup_{\substack{t\in\mathscr{B}\\x\in\mathscr{X}}}\left|L_1^{1/2}\nu_{L_1}(t,x)\right|\geq\delta\right]\leq\frac{\overline{D}}{\delta},\quad Pr\left[\sup_{\substack{t\in\mathscr{B}\\x\in\mathscr{X}}}\left|L_2^{1/2}\nu_{L_2}(t,x)\right|\geq\delta\right]\leq\frac{\overline{D}}{\delta}.$$

Next, note that we can express

$$\widehat{\phi}_k(t|x)-\phi_k(t|x)=\frac{1}{h^c_{x,L_2}}\cdot\nu_{L_2}(t,x)-\frac{1}{h^c_{x,L_1}}+\xi^{F_{0,k}}_{L_2}(t,x)-\xi^{G_1}_{L_1}(t,x),$$

where by the uniform asymptotic properties of these remainder terms, there exists a constant $\overline{C}$ such that, for any $\delta>0$,

$$Pr\left[\sup_{\substack{t\in\mathscr{B}\\x\in\mathscr{X}}}\left|\xi^{G_1}_{L_1}(t,x)\right|>\delta\right]\leq\frac{\overline{C}}{\delta\cdot L_1^{1/2+\epsilon}},\quad\text{and}\quad Pr\left[\sup_{\substack{t\in\mathscr{B}\\x\in\mathscr{X}}}\left|\xi^{F_{o,k}}_{L_2}(t,x)\right|>\delta\right]\leq\frac{\overline{C}}{\delta\cdot L_2^{1/2+\epsilon}},$$

for some $\epsilon>0$. Combining these results, we have

$$Pr\left[\sup_{\substack{t\in\mathscr{B}\\x\in\mathscr{X}}}\left|\widehat{\phi}_k(t|x)-\phi_k(t|x)\right|>\gamma_{L_1}\right]\leq Pr\left[\sup_{\substack{t\in\mathscr{B}\\x\in\mathscr{X}}}\left|\frac{1}{h^c_{x,L_1}}\nu_{L_1}(t,x)\right|>\frac{\gamma_{L_1}}{4}\right]+$$

$$Pr\left[\sup_{\substack{t\in\mathscr{B}\\x\in\mathscr{X}}}\left|\frac{1}{h^c_{x,L_2}}\nu_{L_2}(t,x)\right|>\frac{\gamma_{L_1}}{4}\right]+Pr\left[\sup_{\substack{t\in\mathscr{B}\\x\in\mathscr{X}}}\left|\xi^{G_1}_{L_1}(t,x)\right|>\frac{\gamma_{L_1}}{4}\right]+Pr\left[\sup_{\substack{t\in\mathscr{B}\\x\in\mathscr{X}}}\left|\xi^{F_{0,k}}_{L_2}(t,x)\right|>\frac{\gamma_{L_1}}{4}\right]$$

$$\leq\frac{4\overline{D}}{\gamma_{L_1}h^c_{x,L_1}L_1^{1/2}}+\frac{4\overline{D}}{\gamma_{L_1}h^c_{x,L_2}L_2^{1/2}}+\frac{4\overline{C}}{\gamma_{L_1}L_1^{1/2+\epsilon}}+\frac{4\overline{C}}{\gamma_{L_1}L_2^{1/2+\epsilon}}\longrightarrow 0\ \text{ as }L_1\to\infty,\ L_2\to\infty.$$

Where the last result follows from the bandwidth convergence conditions in Assumption 3.9. Now let us go back to the indicator function $1\left[\sup_{\substack{t\in\mathscr{B}\\x\in\mathscr{X}}}\left|\widehat{\phi}_k(t|x)-\phi_k(t|x)\right|\geq\gamma_{L_1}\right]$ which appears in (A.27). Fix any $\Delta>0$ and $\delta>0$. Then,

$$Pr\left[L_1^{\Delta}1\left[\sup_{\substack{t\in\mathscr{B}\\x\in\mathscr{X}}}\left|\widehat{\phi}_k(t|x)-\phi_k(t|x)\right|\geq\gamma_{L_1}\right]>\delta\right]\leq Pr\left[\sup_{\substack{t\in\mathscr{B}\\x\in\mathscr{X}}}\left|\widehat{\phi}_k(t|x)-\phi_k(t|x)\right|>\gamma_{L_1}\right]\longrightarrow 0.$$

Therefore,

$$1\left[\sup_{\substack{t\in\mathscr{B}\\x\in\mathscr{X}}}\left|\widehat{\phi}_k(t|x)-\phi_k(t|x)\right|\geq\gamma_{L_1}\right]=o_p\left(L_1^{-\Delta}\right)\quad\text{for any }\Delta>0.\tag{A.28}$$

Lastly, the above conditions also imply

$$\sup_{\substack{t\in\mathscr{B}\\x\in\mathscr{X}}}\left|\widehat{\phi}_k(t|x)-\phi_k(t|x)\right|=O_p\left(\frac{1}{L_1^{1/2}h_{x,L_1}^c}\right)+O_p\left(\frac{1}{L_2^{1/2}h_{x,L_2}^c}\right). \tag{A.29}$$

Next, define the following empirical process $\nu_{L_1}^{\phi_k}(\cdot)$ indexed over $\mathscr{B}$,

$$\nu_{L_1}^{\phi_k}(t)=\frac{1}{L_1}\sum_{i=1}^{L_1}\Big(1\left[-2\gamma_{L_1}\leq\phi_k(t|X_i)<0\right]\cdot W_{\mathscr{X}}(X_i)-E\left[1\left[-2\gamma_{L_1}\leq\phi_k(t|X_i)<0\right]\cdot W_{\mathscr{X}}(X_i)\right]\Big).$$

By Assumption 3.8, Lemma 2.12 in Pakes and Pollard (1989) and the Main Corollary in Sherman (1994), this process satisfies $\sup_{t\in\mathscr{B}}\left|\nu_{L_1}^{\phi_k}(t)\right|=O_p\left(L_1^{-1/2}\right)$. Combined with Assumption 3.7, this obtains

$$\begin{aligned}&\sup_{t\in\mathscr{B}}\frac{1}{L_1}\sum_{i=1}^{L_1}1\left[-2\gamma_{L_1}\leq\phi_k(t|X_i)<0\right]\cdot W_{\mathscr{X}}(X_i)\\&\quad\leq\sup_{t\in\mathscr{B}}E\left[1\left[-2\gamma_{L_1}\leq\phi_k(t|X_i)<0\right]\cdot W_{\mathscr{X}}(X_i)\right]+\sup_{t\in\mathscr{B}}\left|\nu_{L_1}^{\phi_k}(t)\right|=O\left(\gamma_{L_1}\right)+O_p\left(L_1^{-1/2}\right)=O_p\left(\gamma_{L_1}\right).\end{aligned} \tag{A.30}$$

Plugging in the results from (A.28), (A.29) and (A.30) into (A.26) and (A.27), for any $\Delta>0$ we obtain

$$\begin{aligned}\frac{1}{L_1}\sum_{i=1}^{L_1}\widehat{\Lambda}_{k,1}(X_i)W_{\mathscr{X}}(X_i)&=\frac{1}{L_1}\sum_{i=1}^{L_1}\int_{t\in\mathscr{B}(X_i)}\widehat{\phi}_k(t|X_i)\cdot 1\left[\phi_k(t|X_i)\geq 0\right]dQ(t|X_i)W_{\mathscr{X}}(X_i)+\widehat{\vartheta}_{k,1},\\ \text{where}\quad\widehat{\vartheta}_{k,1}&=O_p\left(\gamma_{L_1}^2\right)+o_p\left(L_1^{-\Delta}\right)+O_p\left(\frac{\gamma_{L_1}}{L_1^{1/2}h_{x,L_1}^c}\right)+O_p\left(\frac{\gamma_{L_1}}{L_2^{1/2}h_{x,L_2}^c}\right)\\&=O_p\left(L_1^{-1/2-\epsilon}\right)+O_p\left(L_2^{-1/2-\epsilon}\right)\quad\text{for some }\epsilon>0.\end{aligned} \tag{A.31}$$

The same steps can be used to show the equivalent result for $\widehat{\Lambda}_{k,1}$. Namely,

$$\begin{aligned}&\frac{1}{L_2}\sum_{i=1}^{L_2}\widehat{\Lambda}_{k,2}(X_i)W_{\mathscr{X}}(X_i)=\frac{1}{L_2}\sum_{i=1}^{L_2}\int_{t\in\mathscr{B}(X_i)}\widehat{\phi}_k(t|X_i)\cdot 1\left[\phi_k(t|X_i)\geq 0\right]dQ(t|X_i)W_{\mathscr{X}}(X_i)+\widehat{\vartheta}_{k,2},\\&\text{where}\quad\widehat{\vartheta}_{k,2}=O_p\left(\gamma_{L_2}^2\right)+o_p\left(L_2^{-\Delta}\right)+O_p\left(\frac{\gamma_{L_2}}{L_1^{1/2}h_{x,L_1}^c}\right)+O_p\left(\frac{\gamma_{L_2}}{L_2^{1/2}h_{x,L_2}^c}\right)\\&\qquad=O_p\left(L_1^{-1/2-\epsilon}\right)+O_p\left(L_2^{-1/2-\epsilon}\right)\quad\text{for some }\epsilon>0.\end{aligned} \tag{A.32}$$

Recall that $\Lambda_k(x)=\int_{x\in\mathscr{B}(x)}\left(\phi_k(t|x)\right)_+dQ(t|x)$. Combining (A.25) and (A.31),

$$
\begin{aligned}
&\frac{1}{L_1}\sum_{i=1}^{L_1}\widehat{\Lambda}_{k,1}(X_i)W_{\mathscr{X}}(X_i) = E_{1,X}\left[\Lambda_k(X_i)W_{\mathscr{X}}(X_i)\right] + \frac{1}{L_1}\sum_{i=1}^{L_1}\Big(\Lambda_k(X_i)W_{\mathscr{X}}(X_i) - E_{1,X}\left[\Lambda_k(X_i)W_{\mathscr{X}}(X_i)\right]\Big)\\
&+\frac{1}{L_1L_2h_{x,L_2}^c}\sum_{i=1}^{L_1}\sum_{\ell=1}^{L_2}\int_{t\in\mathscr{B}(X_i)}\varphi^{F_{0,k}}\left(P_\ell, X_\ell, t, X_i; h_{p,L_2}, h_{x,L_2}\right)1\left[\phi_k(t|X_i)\geq 0\right]dQ(t|X_i)\cdot W_{\mathscr{X}}(X_i)\\
&-\frac{1}{L_1^2h_{x,L_1}^c}\sum_{i=1}^{L_1}\sum_{m=1}^{L_1}\int_{t\in\mathscr{B}(X_i)}\varphi^{G_1}\left(\boldsymbol{b}_m, X_m, t, X_i; h_{b,L_1}, h_{x,L_1}\right)1\left[\phi_k(t|X_i)\geq 0\right]dQ(t|X_i)W_{\mathscr{X}}(X_i) + \widehat{\varpi}_{k,1},\\
&\quad\text{where}\quad \widehat{\varpi}_{k,1} = O_p\left(L_1^{-1/2-\epsilon}\right) + O_p\left(L_2^{-1/2-\epsilon}\right)\quad\text{for some }\epsilon>0.
\end{aligned}\tag{A.33}
$$

Similarly, combining (A.25) and (A.32),

$$
\begin{aligned}
&\frac{1}{L_2}\sum_{i=1}^{L_2}\widehat{\Lambda}_{k,2}(X_i)W_{\mathscr{X}}(X_i) = E_{2,X}\left[\Lambda_k(X_i)W_{\mathscr{X}}(X_i)\right] + \frac{1}{L_2}\sum_{i=1}^{L_2}\Big(\Lambda_k(X_i)W_{\mathscr{X}}(X_i) - E_{2,X}\left[\Lambda_k(X_i)W_{\mathscr{X}}(X_i)\right]\Big)\\
&+\frac{1}{L_2^2h_{x,L_2}^c}\sum_{i=1}^{L_2}\sum_{\ell=1}^{L_2}\int_{t\in\mathscr{B}(X_i)}\varphi^{F_{0,k}}\left(P_\ell, X_\ell, t, X_i; h_{p,L_2}, h_{x,L_2}\right)1\left[\phi_k(t|X_i)\geq 0\right]dQ(t|X_i)\cdot W_{\mathscr{X}}(X_i)\\
&-\frac{1}{L_2L_1h_{x,L_1}^c}\sum_{i=1}^{L_2}\sum_{m=1}^{L_1}\int_{t\in\mathscr{B}(X_i)}\varphi^{G_1}\left(\boldsymbol{b}_m, X_m, t, X_i; h_{b,L_1}, h_{x,L_1}\right)1\left[\phi_k(t|X_i)\geq 0\right]dQ(t|X_i)W_{\mathscr{X}}(X_i) + \widehat{\varpi}_{k,2},\\
&\quad\text{where}\quad \widehat{\varpi}_{k,2} = O_p\left(L_1^{-1/2-\epsilon}\right) + O_p\left(L_2^{-1/2-\epsilon}\right)\quad\text{for some }\epsilon>0.
\end{aligned}\tag{A.34}
$$

The result in Theorem 3.1 will follow from the Hoeffding decompositions of the U-statistics (see Lemma 5.1.A in Serfling (1980)) that appear on the right-hand sides of Equations (A.33) and (A.34). I will examine each term at a time. Let

$$
U_{L_1}^a = \frac{1}{L_1^2h_{x,L_1}^c}\sum_{i=1}^{L_1}\sum_{m=1}^{L_1}\int_{t\in\mathscr{B}(X_i)}\varphi^{G_1}\left(\boldsymbol{b}_m, X_m, t, X_i; h_{b,L_1}, h_{x,L_1}\right)1\left[\phi_k(t|X_i)\geq 0\right]dQ(t|X_i)W_{\mathscr{X}}(X_i)
$$

Note first that $E_{1,(\boldsymbol{b},X)}\left[\int_{t\in\mathscr{B}(X_i)}\varphi^{G_1}\left(\boldsymbol{b}_m, X_m, t, X_i; h_{b,L_1}, h_{x,L_1}\right)\Big|X_i\right] = 0$ (see the paragraph preceding Equation (A.1)). Fix $\boldsymbol{b}$ and x and let

$$
\begin{aligned}
&\Xi_1^{G_1}\left(\boldsymbol{b}, x, X^c; h_{b,L_1}, h_{x,L_1}\right) = E_{1,X^d|X^c}\left[\int_{t\in\mathscr{B}(X)}\varphi^{G_1}\left(\boldsymbol{b}, x, t, X; h_{b,L_1}, h_{x,L_1}\right)1\left[\phi_k(t|X)\geq 0\right]dQ(t|X)W_{\mathscr{X}}(X)\Big|X^c\right],\\
&\lambda_1^{G_1}(\boldsymbol{b}_i, X_i; h_{b,L_1}, h_{x,L_1}) = \int_{\psi\in[-S,S]^c}\Xi_1^{G_1}\left(\boldsymbol{b}_i, X_i, h_{x,L_1}\psi + X_i^c; h_{b,L_1}, h_{x,L_1}\right)f_{1,X^c}(h_{x,L_1}\psi + X_i^c)d\psi.
\end{aligned}
$$

Notice that $E_{1,(\boldsymbol{b},X)}\left[\lambda_1^{G_1}(\boldsymbol{b}_i, X_i; h_{b,L_1}, h_{x,L_1})\right] = 0$ since, by iterated expectations,

$$E_{1,(\boldsymbol{b},X)}\left[\lambda_1^{G_1}(\boldsymbol{b}_i,X_i;h_{b,L_1},h_{x,L_1})\right]=$$

$$E_{1,X}\left[\underbrace{E_{1,(\boldsymbol{b},X)}\left[\int_{t\in\mathscr{B}(X_i)}\varphi^{G_1}\left(\boldsymbol{b}_m,X_m,t,X_i;h_{b,L_1},h_{x,L_1}\right)\Big|X_i\right]}_{=0}1\left[\phi_k(t|X_i)\geq 0\right]dQ(t|X_i)W_{\mathscr{X}}(X_i)\right]=0.$$

The Hoeffding decomposition of $U_{L_1}^a$ yields the following result,

$$U_{L_1}^a=\frac{1}{L_1}\sum_{i=1}^{L_1}\lambda_1^{G_1}(\boldsymbol{b}_i,X_i;h_{b,L_1},h_{x,L_1})+\widehat{\nu}^a,\ \text{ where }\widehat{\nu}^a=O_p\left(\frac{1}{L_1h_{x,L_1}^c}\right)=O_p\left(L_1^{-1/2-\epsilon}\right)\quad\text{for some }\epsilon>0.\tag{A.35}$$

Where the last result follows from the bandwidth convergence conditions in Assumption 3.9. Next, let

$$U_{L_2}^b=\frac{1}{L_2^2h_{x,L_2}^c}\sum_{i=1}^{L_2}\sum_{\ell=1}^{L_2}\int_{t\in\mathscr{B}(X_i)}\varphi^{F_{0,k}}\left(P_\ell,X_\ell,t,X_i;h_{p,L_2},h_{x,L_2}\right)1\left[\phi_k(t|X_i)\geq 0\right]dQ(t|X_i)\cdot W_{\mathscr{X}}(X_i).$$

And note (from (A.23)) that $E_{2,(P,X)}\left[\varphi^{F_{0,k}}\left(P_\ell,X_\ell,t,X_i;h_{p,L_2},h_{x,L_2}\right)\big|X_i\right]=0$. Fix p and x and let

$$\Xi_2^{F_{0,k}}\left(p,x,X^c;h_{p,L_2},h_{x,L_2}\right)=E_{2,X^d|X^c}\left[\int_{t\in\mathscr{B}(X)}\varphi^{F_{0,k}}\left(p,x,t,X;h_{p,L_2},h_{x,L_2}\right)1\left[\phi_k(t|X)\geq 0\right]dQ(t|X)W_{\mathscr{X}}(X)\Big|X^c\right.$$

$$\lambda_2^{F_{0,k}}(P_i,X_i;h_{p,L_2},h_{x,L_2})=\int_{\psi\in[-S,S]^c}\Xi_2^{F_{0,k}}\left(P_i,X_i,h_{x,L_2}\psi+X_i^c;h_{p,L_2},h_{x,L_2}\right)f_{2,X^c}(h_{x,L_2}\psi+X_i^c)d\psi$$

Note that $E_{2,(P,X)}\left[\lambda_2^{F_{0,k}}(P_i,X_i;h_{p,L_2},h_{x,L_2})\right]=0$ since, by iterated expectations,

$$E_{2,(P,X)}\left[\lambda_2^{F_{0,k}}(P_i,X_i;h_{p,L_2},h_{x,L_2})\right]=$$

$$E_{2,X}\left[\underbrace{E_{2,(P,X)}\left[\int_{t\in\mathscr{B}(X_i)}\varphi^{F_{0,k}}\left(P_\ell,X_\ell,t,X_i;h_{p,L_2},h_{x,L_2}\right)\Big|X_i\right]}_{=0}1\left[\phi_k(t|X_i)\geq 0\right]dQ(t|X_i)W_{\mathscr{X}}(X_i)\right]$$

The Hoeffding decomposition of $U_{L_2}^b$ yields the following result,

$$U_{L_2}^b=\frac{1}{L_2}\sum_{i=1}^{L_2}\lambda_2^{F_{0,k}}(P_i,X_i;h_{p,L_2},h_{x,L_2})+\widehat{\nu}^b,\ \text{ where }\widehat{\nu}^b=O_p\left(\frac{1}{L_2h_{x,L_2}^c}\right)=O_p\left(L_2^{-1/2-\epsilon}\right)\quad\text{for some }\epsilon>0.\tag{A.36}$$

Once again the last line results from the bandwidth properties described in Assumption 3.9. Next let us analyze the generalized U-statistic[30] (see Section 5.1.3 in Serfling (1980)) ,

$$U^c_{L_1,L_2} = \frac{1}{L_2 L_1 h^c_{x,L_1}} \sum_{i=1}^{L_2} \sum_{m=1}^{L_1} \int_{t\in\mathscr{B}(X_i)} \varphi^{G_1}\left(\boldsymbol{b}_m, X_m, t, X_i; h_{b,L_1}, h_{x,L_1}\right) 1\left[\phi_k(t|X_i) \geq 0\right] dQ(t|X_i) W_{\mathscr{X}}(X_i).$$

Note that $E_{1,(\boldsymbol{b},X)}\left[\varphi^{G_1}\left(\boldsymbol{b}_m, X_m, t, X_i; h_{b,L_1}, h_{x,L_1}\right)|X_i\right] = 0$ (again, see the paragraph preceding Equation (A.1)). Fix $\boldsymbol{b}$ and x and let

$$\Xi^{G_1}_2\left(\boldsymbol{b}, x, X^c, h_{b,L_1}, h_{x,L_1}\right) = E_{2,X^d|X^c}\left[\int_{t\in\mathscr{B}(X)} \varphi^{G_1}\left(\boldsymbol{b}, x, t, X; h_{b,L_1}, h_{x,L_1}\right) 1\left[\phi_k(t|X) \geq 0\right] dQ(t|X) W_{\mathscr{X}}(X) \middle| X^c\right],$$

$$\lambda^{G_1}_2(\boldsymbol{b}_i, X_i; h_{b,L_1}, h_{x,L_1}) = \int_{\psi\in[-S,S]^c} \Xi^{G_1}_2\left(\boldsymbol{b}_i, X_i, h_{x,L_1}\psi + X^c_i, h_{b,L_1}, h_{x,L_1}\right) f_{2,X^c}(h_{x,L_1}\psi + X^c_i) d\psi$$

As in the previous cases, the last functional satisfies $E_{1,(\boldsymbol{b},X)}\left[\lambda^{G_1}_2(\boldsymbol{b}_i, X_i; h_{b,L_1}, h_{x,L_1})\right] = 0$. This can be shown by iterated expectations, since

$$E_{1,(\boldsymbol{b},X)}\left[\lambda^{G_1}_2(\boldsymbol{b}_i, X_i; h_{b,L_1}, h_{x,L_1})\right] =$$
$$E_{2,X}\left[\underbrace{E_{1,(\boldsymbol{b},X)}\left[\int_{t\in\mathscr{B}(X)} \varphi^{G_1}\left(\boldsymbol{b}_\ell, X_\ell, t, X_i, h_{b,L_1}, h_{x,L_1}\right) \middle| X_i\right]}_{=0} 1\left[\phi_k(t|X_i) \geq 0\right] dQ(t|X_i) W_{\mathscr{X}}(X_i)\right] = 0$$

A Hoeffding decomposition result for the generalized U-statistic $U^c_{L_1,L_2}$ results in the following,

$$\begin{aligned} U^c_{L_1,L_2} &= \frac{1}{L_1}\sum_{i=1}^{L_1} \lambda^{G_1}_2(\boldsymbol{b}_i, X_i; h_{b,L_1}, h_{x,L_1}) + \widehat{\nu}^c, \\ &\quad \text{where} \quad \widehat{\nu}^c = O_p\left(\frac{1}{\sqrt{L_1 L_2} h^c_{x,L_1}}\right) = \begin{cases} O_p\left(L_1^{-1/2-\epsilon}\right) \\ O_p\left(L_2^{-1/2-\epsilon}\right) \end{cases} \text{for some } \epsilon > 0. \end{aligned} \tag{A.37}$$

The last result being a consequence of Assumption 3.9. The last term left to simplify the asymptotic properties of expressions (A.32) and (A.33) is the following generalized U-statistic,

$$U^d_{L_1,L_2} = \frac{1}{L_1 L_2 h^c_{x,L_2}} \sum_{i=1}^{L_1} \sum_{\ell=1}^{L_2} \int_{t\in\mathscr{B}(X_i)} \varphi^{F_{0,k}}\left(P_\ell, X_\ell, t, X_i; h_{p,L_2}, h_{x,L_2}\right) 1\left[\phi_k(t|X_i) \geq 0\right] dQ(t|X_i) \cdot W_{\mathscr{X}}(X_i)$$

[30] The extension of Hoeffding's decomposition results (Hoeffding (1948)), from one sample U-statistics to the general case of combining data from k samples, dates back to Lehmann (1951) and Dwas (1956). See Section 5.1.3 in Serfling (1980).

As it was the case with the statistics $U^a_{L_1}$, $U^b_{L_2}$ and $U^c_{L_1,L_2}$, from Equation (A.23) we have in this case that $E_{2,X^d|X^c}\left[\varphi^{F_{0,k}}\left(P_\ell, X_\ell, t, X_i; h_{p,L_2}, h_{x,L_2}\right)\big|X_i\right] = 0$. Fix p, x and let

$$\Xi_1^{F_{0,k}}(p,x,X^c;h_{p,L_2},h_{x,L_2}) = E_{1,X^d|X^c}\left[\int_{t\in\mathscr{B}(X)} \varphi^{F_{0,k}}(p,x,t,X;h_{p,L_2},h_{x,L_2})1\left[\phi_k(t|X)\geq 0\right]dQ(t|X)W_{\mathscr{X}}(X)\middle| X^c\right]$$

$$\lambda_1^{F_{0,k}}(P_i,X_i;h_{p,L_2},h_{x,L_2}) = \int_{\psi\in[-S,S]^c} \Xi_1^{F_{0,k}}(P_i,X_i,h_{x,L_2}\psi + X_i^c;h_{p,L_2},h_{x,L_2})f_{1,X^c}(h_{x,L_2}\psi + X_i^c)d\psi$$

The last functional satisfies $E_{2,(P,X)}\left[\lambda_1^{F_{0,k}}(P_i,X_i;h_{p,L_2},h_{x,L_2})\right] = 0$ since, by iterated expectations,

$$E_{2,(P,X)}\left[\lambda_1^{F_{0,k}}(P_i,X_i;h_{p,L_2},h_{x,L_2})\right] =$$
$$E_{1,X}\left[\int_{t\in\mathscr{B}(X_\ell)} \underbrace{E_{2,(P,X)}\left[\varphi^{F_{0,k}}(P_i,X_i,t,X_\ell;h_{p,L_2},h_{x,L_2})\big|X_\ell\right]}_{=0} 1\left[\phi_k(t|X_\ell)\geq 0\right]dQ(t|X_\ell)W_{\mathscr{X}}(X_\ell)\right] = 0.$$

The Hoeffding decomposition expression for $U^d_{L_1,L_2}$ is,

$$U^d_{L_1,L_2} = \frac{1}{L_2}\sum_{i=1}^{L_2}\lambda_1^{F_{0,k}}(P_i,X_i;h_{p,L_2},h_{x,L_2}) + \widehat{\nu}^d,$$
$$\text{where}\quad \widehat{\nu}^d = O_p\left(\frac{1}{\sqrt{L_1L_2}h^c_{x,L_2}}\right) = \begin{cases} O_p\left(L_1^{-1/2-\epsilon}\right) \\ O_p\left(L_2^{-1/2-\epsilon}\right)\end{cases} \text{for some } \epsilon > 0. \tag{A.38}$$

Combining the previous results with (A.33),

$$\frac{1}{L_1}\sum_{i=1}^{L_1}\widehat{\Lambda}_{k,1}(X_i)W_{\mathscr{X}}(X_i) = E_{1,X}\left[\Lambda_k(X_i)W_{\mathscr{X}}(X_i)\right] + \frac{1}{L_1}\sum_{i=1}^{L_1}\Big(\Lambda_k(X_i)W_{\mathscr{X}}(X_i) - E_{1,X}\left[\Lambda_k(X_i)W_{\mathscr{X}}(X_i)\right]\Big)$$
$$+\frac{1}{L_2}\sum_{i=1}^{L_2}\lambda_1^{F_{0,k}}(P_i,X_i;h_{p,L_2},h_{x,L_2}) - \frac{1}{L_1}\sum_{i=1}^{L_1}\lambda_1^{G_1}(\boldsymbol{b}_i,X_i;h_{b,L_1},h_{x,L_1}) + \widehat{\nu}_1,$$
$$\text{where}\quad \widehat{\nu}_1 = O_p\left(L_1^{-1/2-\epsilon}\right) + O_p\left(L_1^{-1/2-\epsilon}\right)\quad \text{for some } \epsilon > 0. \tag{A.39}$$

And Equation (A.34) becomes,

$$\frac{1}{L_2}\sum_{i=1}^{L_2}\widehat{\Lambda}_{k,2}(X_i)W_{\mathscr{X}}(X_i) = E_{2,X}\left[\Lambda_k(X_i)W_{\mathscr{X}}(X_i)\right] + \frac{1}{L_2}\sum_{i=1}^{L_2}\Big(\Lambda_k(X_i)W_{\mathscr{X}}(X_i) - E_{2,X}\left[\Lambda_k(X_i)W_{\mathscr{X}}(X_i)\right]\Big)$$
$$+\frac{1}{L_2}\sum_{i=1}^{L_2}\lambda_2^{F_{0,k}}(P_i,X_i;h_{p,L_2},h_{x,L_2}) - \frac{1}{L_1}\sum_{i=1}^{L_1}\lambda_2^{G_1}(\boldsymbol{b}_i,X_i;h_{b,L_1},h_{x,L_1}) + \widehat{\nu}_2, \tag{A.40}$$

$$\text{where} \quad \widehat{\nu}_2 = O_p\left(L_1^{-1/2-\epsilon}\right) + O_p\left(L_1^{-1/2-\epsilon}\right) \quad \text{for some } \epsilon > 0.$$

Let

$$\begin{aligned}
\psi_{1,k}(\boldsymbol{b}_j, X_j; h_{b,L_1}, h_{x,L_1}) &= (\Lambda_k(X_j)W_{\mathscr{X}}(X_j) - E_{1,X}\left[\Lambda_k(X_j)W_{\mathscr{X}}(X_j)\right]) \\
&\quad - \lambda_1^{G_1}(\boldsymbol{b}_j, X_j; h_{b,L_1}, h_{x,L_1}) - \lambda_2^{G_1}(\boldsymbol{b}_j, X_j; h_{b,L_1}, h_{x,L_1}), \\
\psi_{2,k}(P_i, X_i; h_{p,L_2}, h_{x,L_2}) &= (\Lambda_k(X_i)W_{\mathscr{X}}(X_i) - E\left[\Lambda_k(X_i)W_{\mathscr{X}}(X_i)\right]) \\
&\quad + \lambda_1^{F_{0,k}}(P_i, X_i; h_{p,L_2}, h_{x,L_2}) + \lambda_2^{F_{0,k}}(P_i, X_i; h_{p,L_2}, h_{x,L_2}).
\end{aligned} \tag{A.41}$$

Let $L \equiv L_1 + L_2$. Then, for some $\Delta > \frac{1}{2}$,

$$\begin{aligned}
\widehat{\mathcal{T}}_k &= \mathcal{T}_k + \frac{1}{L_1}\sum_{j=1}^{L_1}\psi_{1,k}(\boldsymbol{b}_j, X_j; h_{b,L_1}, h_{x,L_1}) + \frac{1}{L_2}\sum_{i=1}^{L_2}\psi_{2,k}(P_i, X_i; h_{p,L_2}, h_{x,L_2}) + o_p\left(L_1^{-\Delta}\right) + o_p\left(L_2^{-\Delta}\right) \\
\widehat{\mathcal{T}}_k &= \mathcal{T}_k + \frac{1}{L_1}\sum_{j=1}^{L_1}\psi_{1,k}(\boldsymbol{b}_j, X_j; h_{b,L_1}, h_{x,L_1}) + \frac{1}{L_2}\sum_{i=1}^{L_2}\psi_{2,k}(P_i, X_i; h_{p,L_2}, h_{x,L_2}) + o_p\left(L^{-\Delta}\right).
\end{aligned} \tag{A.42}$$

This proves Theorem 3.1. ■