ESSAYS ON AUCTION INFERENCE WITH NON-EQUILIBRIUM BELIEFS

A Dissertation in Economics
by
Lidia Kosenkova

© 2018 Lidia Kosenkova

Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

August 2018
The dissertation of Lidia Kosenkova was reviewed and approved* by the following:

Andrés Aradillas-López
Associate Professor of Economics
Dissertation Advisor, Chair of Committee

Joris Pinkse
Professor of Economics

Peter Newberry
Assistant Professor of Economics

Bing Li
Professor of Statistics

Barry W. Ickes
Professor of Economics
Department Head

*Signatures are on file in the Graduate School.
Abstract

In this dissertation I study bidding behavior in first-price sealed bid auctions with risk-neutral bidders. Instead of assuming that bids and beliefs correspond to a Bayesian Nash equilibrium (BNE), I only assume that they are consistent with $k$ steps of iterated elimination of dominated strategies ($k$—rationalizability). The focus of my dissertation is to provide econometric tests for whether $k$ is finite and to identify the largest value of $k$ that is consistent with the data. This is important because rejecting any finite $k$ would immediately rule out BNE and (full) rationalizability and it allows to quantify deviations from (fully) rationalizable behavior 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 tests to USFS timber auction data. The results show that values of $k$ as low as 2 can be rejected in some auctions. Counterfactual exercises allow me to quantify the loss in expected payoff derived from the presence of incorrect beliefs.
# Table of Contents

List of Figures vii

List of Tables viii

Acknowledgments ix

Chapter 1

Nonparametric inference in asymmetric first-price auctions with k-rationalizable beliefs 1

1.1 Introduction ............................................................... 1

1.2 A Model of k-rationalizable Bidding ................................. 6
  1.2.1 Rationalizable bids ................................................. 7
  1.2.1.1 Basic setup ..................................................... 7
  1.2.1.2 Beliefs and best responses .................................... 7
  1.2.1.3 Rationalizable bids for a given upper bound on oppo-
  nents' bidding functions ............................................ 8
  1.2.1.4 An iterative characterization of rationalizable bids .... 9
  1.2.2 Behavior and stylized facts consistent with k-rationalizable
  bidding ......................................................................... 11
  1.2.3 Extensions to asymmetric bidders ................................. 14

1.3 An Econometric test for k-rationalizability ......................... 16
  1.3.1 k-rationalizability as a stochastic dominance restriction .... 16
  1.3.2 A nonparametric test for k-rationalizability with auxiliary data
    from ascending auctions ........................................... 17
  1.3.2.1 An econometric test .......................................... 21
  1.3.2.2 Nonparametric estimators from the sample of ascend-
    ing auctions ........................................................ 24
  1.3.2.3 Nonparametric estimators from the sample of first-
    price auctions ...................................................... 26
  1.3.2.4 Estimation of $T_k$ ............................................. 26
1.3.2.5 A test based on Theorem 1.3.1 .................................................. 31
1.4 Extensions .................................................................................. 34
1.4.1 A nonparametric test when the distribution of private values is only partially identified. .................................................. 34
1.4.2 A semi-parametric model of first-price auctions with independent private values without auxiliary data .................................. 37
1.4.3 k-rationalizability with ambiguity ............................................. 41
1.4.4 A semi-parametric model of first-price auctions with interdependent values and affiliated signals ................................................. 44
1.4.5 Risk-aversion ........................................................................... 46
1.4.6 Inference when there is collusion in ascending auctions .......... 50
1.4.7 Testing when the assumption $F_{0,1} = F_{0,2}$ is misspecified .......... 50
1.4.7.1 Endogenous participation .................................................... 51
1.5 Concluding remarks .................................................................... 52

Chapter 2

Testing $k$–rationalizability in USFS timber auctions 54
2.1 Introduction .................................................................................. 54
2.2 Data description .......................................................................... 55
  2.2.1 Choosing $X$, the vector of conditioning variables ................. 58
2.3 Collusion ...................................................................................... 60
2.4 Implementation of the test ............................................................ 62
2.5 Tests for risk-neutral BNE ............................................................ 63
  2.5.1 A test for no overbidding above risk-neutral BNE ................. 64
  2.5.2 A test for risk-averse BNE ...................................................... 65
  2.5.3 Two-sided BNE test ................................................................. 65
  2.5.4 Results for BNE tests .............................................................. 65
2.6 A counterfactual exercise to estimate the economic losses from incorrect beliefs ................................................................. 66
  2.6.1 Three measures of expected payoffs ...................................... 67
  2.6.2 Results from comparing current expected payoffs against counterfactuals ................................................................. 69
2.7 Concluding remarks ..................................................................... 71

Chapter 3

Testing functional inequalities conditional on an estimated index 75
3.1 Introduction .................................................................................. 75
3.2 Model .......................................................................................... 76
3.3 Testing procedure ......................................................................... 77
  3.3.1 An estimator for $R(t)$ .............................................................. 78
List of Figures

1.1 Comparison of nonparametrically estimated bidding functions, BNE bidding functions and $k$–rationalizable bounds for $k = 2$.

1.2 Graphical illustration of conditions (1.15), (1.16) and (1.17) in the proof of Result 1.4.1.

1.3 Truncated exponential family example. Distribution, lower envelope and $k$–rationalizable bounds (for $k = 6$).

1.4 $k$–rationalizable bounds with risk-aversion (for $k = 6$). An illustration with CRRA utility function and $U[0, 1]$ valuations.

2.1 Distribution of bidder size. A comparison of first-price and ascending auctions.

2.2 Density of the reserve price in ascending and sealed-bid auctions with $n=2$ players.

2.3 Expected payoffs conditional on reserve price.

2.4 Expected payoffs conditional on reserve price, with 95% bootstrap point-wise confidence bands.
List of Tables

2.1 Summary statistics .................................................. 56
2.2 Choice of auction type (1=sealed-bid auction) .................. 57
2.3 Determinants of transaction price. .............................. 59
2.4 Results from the tests for \( k \)-rationalizability ............ 63
2.5 Results of BNE tests in auctions with bidders of the same type. 66
2.6 Counterfactual results: Difference in large bidders’ expected payoffs (in 2010 real dollars). .......................... 70
2.7 Counterfactual results: Percentage difference in large bidders’ expected payoffs. .............................. 74
2.8 Counterfactual results: Evidence of overbidding. Summary statistics for \( b_{\text{benchmark}}(\cdot) \) and \( b_{\text{boundary}}(\cdot) \) (in % terms) in auctions won by large bidders. 74

B.1 Collusion analysis based on bid levels: Results from the test \( H_0 : \theta^c = \theta^{nc} \) where the set of potential colluders is defined as those bidders in the top \( D^{th} \) quantile in terms of high-valued auctions won during the time period. ................................................. 113

B.2 Collusion analysis based on bid rankings: Testing \( H_0 : \) high-bid and low-bid rankings are generated by the same bidding rule. Set of potential colluders is defined as those bidders in the top \( D^{th} \) quantile in terms of high-valued auctions won during the time period. ................................. 115
Acknowledgments

I gratefully acknowledge the supervision of Andrés Aradillas-López as well as useful comments from Marc Henry, Joris Pinkse, Robert Marshall, Gaurab Aryal, Christian Bontemps, Michael Gechter, Keisuke Hirano, Sung Jae Jun, Nail Kashaev, Vijay Krishna, Rohit Lamba, Salvador Navarro, Bruno Salcedo and Karl Schurter
To my parents: Victor Kosenkov and Galina Kosenkova
Chapter 1
Nonparametric inference in asymmetric first-price auctions with k-rationalizable beliefs

1.1 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 (expected profit-maximizing) bidding strategies depend crucially on bidders’ beliefs. While having incorrect beliefs may result in significant losses in expected-payoff, simply imposing the assumption of correct beliefs—as is the case in Bayesian-Nash equilibrium (BNE)—may be difficult to justify in many real-world applications and it can lead to incorrect econometric inference and predictions. Instead of assuming that beliefs are correct, I rely on a more general family of behavioral models which, while encompassing BNE and rationalizability as special cases, is rich enough to include a wider range of bidding behavior.

I consider the solution concept that assumes only that beliefs (and bidding strategies) are consistent with ‘k’ steps of iterated elimination of dominated strategies, but beliefs can be incorrect\(^1\) otherwise. This is the notion of k—rationalizability. I focus on econometric tests for whether k is finite and to identify the largest value of k that is consistent with the data. If bidding behavior is consistent only with a

\(^1\)Incorrect beliefs refer to subjective probability distributions that do not correspond to the true distributions of opponents’ bidding strategies.
finite $k$, I reject BNE and rationalizability as the true underlying behavioral models. Furthermore, identifying the largest value of $k$ that is consistent with the data allows me to quantify the deviation from rationalizable behavior and to potentially improve counterfactual predictions. The results of my tests in the empirical illustration hint at a particular pattern of incorrect beliefs. Based on these results, I perform a counterfactual analysis to quantify the loss in expected payoffs derived from this misalignment of beliefs. I show that in the presence of auxiliary data from ascending auctions, an econometric test for $k$—rationalizability can be fully nonparametric, while in the absence of such data I can still conduct parametric inference for $k$.

The setting of my paper is first-price sealed-bid auctions for a single object with symmetric risk-neutral bidders. I can allow for asymmetries based on observable characteristics, such as “large” and “small” bidding firms. The model I use to describe $k$—rationalizability allows for common values and affiliation. However, the nonparametric test I present relies on the assumption of Independent Private Values (IPV). Thus, my results allow affiliation in the context of a parametric model, while nonparametric tests and results assume IPV and the presence of auxiliary data from ascending auctions.

Foundational works on the econometric modelling of first-price sealed-bid auctions focused on parametric models and can be traced back to Paarsch (1992) (private and common values), Laffont et al. (1995) (IPV), Donald and Paarsch (1996) (IPV). Subsequently, nonparametric identification results were obtained, for example, in Guerre et al. (2000) (IPV), Li et al. (2000) (conditionally independent private values), Li et al. (2002) (affiliated private values), Hubbard et al. (2012) (affiliated private values). To obtain identification results, each of these papers (and most of the literature) rely on the assumption of Bayesian Nash equilibrium (BNE) behavior. The assumption of equilibrium behavior has also been central for identification results obtained in other auction formats under different paradigms (Hendricks et al. (2003), Athey and Haile (2002), Lu and Perrigne (2008), Athey et al. (2011), Aradillas-López et al. (2013)).

Explicit in the definition of BNE is the assumption that players are expected-payoff maximizers and hold correct beliefs about their opponents’ strategies. Experimental economists have identified deviations from BNE in auctions (Crawford and Iriberri (2007), Kagel and Levin (2014)), and have modeled these violations through cognitive hierarchy or “level-k thinking” models (Crawford and Iriberri (2007), Gillen (2009),
Level-k models assume a hierarchical structure of beliefs based on a naïve $L_0$-type with nonstrategic bidding behavior. Examples of the most common assumptions on the naïve type are that a player of $L_0$-type bids uniformly between zero and her expected value or that he bids her expected value. $L_0$-type’s behavior is assumed to be a common prior. From here, an $L_1$-type of bidder is defined as one who best-responds to the belief that everybody else is $L_0$, $L_2$ best-responds to $L_1$, etc., so that $L_k$-types believe that everybody else is $L(k-1)$. Even if the exact beliefs of the $L_0$ type are not specified (as can be done, under certain conditions; see An (2017)), level-k models have a very rigid structure in terms of how beliefs (and best-responses) are selected: An $L_k$-type always assumes that everybody else is $L(k-1)$ and, in particular, bidders’ beliefs are independent of their values. This cognitive hierarchy gives point prediction of the bidding behavior for each level $k$. In general, level-k models also fail to include BNE as a special case. Furthermore, different specifications of bidding strategy for the $L_0$-type result in different hierarchies of bidders’ beliefs of higher types.

In contrast to level-k cognitive hierarchy models, I do inference based solely on the implications of $k$ steps of iterated elimination of dominated strategies without any assumptions on belief selection\(^2\) or the additional restrictions on “types”. The solution concept I use encompasses rationalizability and BNE as special cases. It also allows for “incorrect” beliefs\(^3\) that are not fully rationalizable and can depend on a bidder’s private value. To achieve this, I rely on the theory of rationalizable bidding in first-price auctions developed by Battigalli and Siniscalchi (2003). They characterize the properties of all bids that can be rationalized as best-responses to beliefs consistent with $k$ steps of iterated deletion of dominated strategies ($k$-rationalizable bids). Under a very mild restriction\(^4\) on the space of allowable beliefs, they show that $k$-rationalizable bids are fully characterized by an upper bound $\overline{B}_k$. The sets of $k$-rationalizable bids are nested, so that $\overline{B}_k \geq \overline{B}_{k+1}$ for all $k$. Thus, the upper bound $\overline{B}_k$ must be satisfied by $k'$-rationalizable bids for all $k' > k$. In particular, they must be satisfied if bids are (fully) rationalizable ($k = \infty$) or if they are BNE bids.

\(^2\)Cognitive hierarchy type models (and other forms of nonequilibrium behavior) are also considered in Kline (2015) in the context of complete-information games. The results presented there also hinge on assumptions about the underlying belief selection mechanism.

\(^3\)I refer to “incorrect beliefs” as those that do not correspond to the actual probability distribution of opponents’ strategies.

\(^4\)The only restriction imposed is that bidders believe that any nonzero bid can win the good with positive probability.
Consequently, a model of $k-$rationalizability based on these bounds encompasses the two most popular solution concepts (BNE and rationalizability) as special cases, but it can also explain many other types of bidding behavior. The goal of my paper is to develop econometric procedures to test if bids observable in the data satisfy the upper bound for $k$-rationalizable bids. If the bounds are violated for a finite $k$, I reject rationalizability (and therefore BNE). The largest $k$, such that the data is consistent with $k$-rationalizable behavior, quantifies the degree of deviation from BNE.

The econometric test is based on inequalities of certain conditional distributions implied by the upper bounds for $k$-rationalizable bids. I show that under the assumption of IPV and the presence of auxiliary data from ascending auctions of the same type of object, the test can be performed nonparametrically. I propose an econometric test based on those inequalities and characterize its asymptotic properties. The test is similar to the one proposed in Aradillas-López et al. (2016), as it adapts to the properties of the so-called contact sets through the use of a tuning parameter. This generates a procedure that has asymptotically pivotal properties, is computationally simple to implement and sidesteps the conservative properties of tests based on “least favorable configurations”.

Among the extensions included, I also discuss how to do inference in a parametric model, which deals with cases where there is no auxiliary data from ascending auctions or when values in first-price auctions are interdependent. In that case, I construct confidence sets for the parameters of the model consistent with the restrictions of $k-$rationalizability (for any pre-specified $k$). An empty confidence set leads to rejection of the hypothesis that all bids in the population are $k$-rationalizable. A similar approach was suggested in Aradillas-Lopez and Tamer (2008).

As an empirical illustration, I use data from the U.S. Forest Service timber tract auctions in Region 6 (Oregon and Washington) 1994-2007. The data contains both ascending and sealed bid types of auctions, and a rich collection of bidder and tract characteristics. Similar to Lu and Perrigne (2008), I use ascending auctions data to nonparametrically estimate the distribution of private values conditional on the observable characteristics of the auction. This allows me to nonparametrically estimate the upper bounds for $k$-rationalizable bids and test whether observable bids in the first-price auctions data satisfy those bounds through a stochastic-dominance test of conditional distributions. This empirical analysis is conducted for asymmetric types of bidders based on firm size (number of employees). Types can have different
distributions of private values and different \( k \)'s depending on the type of bidder they face. The results suggest that values of \( k \) as low as 2 can be rejected when bidders play against bidders of the other type. While larger values of \( k \) are consistent with the observed bids when bidders play face their own type. I reject BNE in all auctions: both with symmetric or asymmetric bidders. The results suggest a misalignment in beliefs in terms of \( k \). Counterfactual exercises quantify the extent to which these incorrect conjectures translate into monetary losses in bidders’ expected profits. I find that correcting beliefs about opponents’ \( k \)-values would lead to a median improvement in expected profits equivalent to eight times the advertised value of the tract. There is evidence that incorrect beliefs lead to widespread overbidding on the most valuable tracts.

The paper proceeds as follows. Section 1.2 describes the model of \( k \)-rationalizable bidding that is used in the paper and characterizes the upper bounds for bidding functions that are consistent with \( k \)-rationalizability. In Section 1.3, I describe econometric tests for \( k \)-rationalizability based on the observable implications of the model. I show that, within an IPV framework, if I have access to auxiliary data from ascending auctions I can conduct a fully nonparametric test. I describe such test and I characterize its asymptotic properties. An empirical illustration for the case of USFS timber auction data is included in Section 2. I classify bidding firms according to the number of workers into “small” and “large” bidders. The results show that values of \( k \) as low as \( k = 2 \) are rejected in asymmetric auctions, while larger values of \( k \) are consistent with the data in symmetric auctions. Thus, there is evidence of incorrect beliefs. Counterfactual analysis quantifies the losses in bidders’ expected profit because of incorrect beliefs about the opponents’ behavior. Outside the IPV framework or without access to ascending auctions data, inference can be carried out parametrically through inequalities of conditional distributions implied by the model. This is outlined in Section 1.4, along with some extensions including the case of collusion in the sample of ascending auctions, bidders’ risk-aversion and unknown distribution of private values (ambiguity). Section 1.5 concludes. Econometric proofs and other supplemental materials are in the Appendix.
1.2 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 be 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 $B_k(\cdot)$ on bidding functions (mappings from bidders’ values to bids). A bidding function $b(\cdot)$ is consistent with $k$-rationalizability if and only if $\forall v b(v) \leq B_k(v)$. These bounds turn out to have useful properties: for each $k$ the upper bound $B_k(\cdot)$ is a strictly increasing function, bounds are monotonically decreasing in $k$, so that $B_k(\cdot) \geq B_{k+1}(\cdot)$ for all $k$, and if I define $B_\infty(v) = \lim_{k \to \infty} B_k(v)$, then $B_\infty(\cdot)$ is a sharp upper bound for the space of rationalizable bids. Since every BNE is rationalizable, each BNE bidding function is below the bound $B_\infty(\cdot)$ 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.
1.2.1 Rationalizable bids

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

1.2.1.1 Basic setup

Consider a single-object first-price auction. There are *n* risk-neutral players with independent private values.\(^5\) Private values are drawn from a commonly known distribution\(^6\) \(F_0\), with compact 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\(^7\) is maintained.

**Assumption 1.2.1** 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.

1.2.1.2 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 \(b_j : [0, \overline{v}] \rightarrow \mathbb{R}_+\). 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 \(\mu_i\) on \(\Delta(\mathcal{B}_{-i})\). I focus on beliefs that assign probability zero to ties. The expected payoff of bidding \(b\) conditional on the private value \(v_i\) and a given belief \(\mu_i \in \Delta(\mathcal{B}_{-i})\) is

\[
\pi(b, v_i; \mu_i) \equiv (v_i - b) \int_{B_{-i}} P[b_{-i} \leq b] \mu_i(db_{-i}) \equiv (v_i - b) \mathbb{P}[b_{-i} \leq b | \mu_i].
\] (1.1)

---

\(^5\)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 1.3.2 relies on the IPV assumption. The parametric approach outlined in Section 1.4.4 can allow for common values under affiliation.

\(^6\)I extend the results to a setting with asymmetric bidders in Section 1.2.3.

\(^7\)In the econometric inference section I impose further smoothness restrictions on \(F_0\).

\(^8\)I focus throughout on *interim* rationalizability, that is, bids are rationalizable for \(i\) conditional on observing \(v_i\).
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 \( \mu_i \).

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 1.2.2** Let
\[ \Delta^+(B_{-i}) = \left\{ \mu \in \Delta(B_{-i}): \forall b > 0, \int_{B_{-i}} P_{v_{-i}} [b_{-i}(v_{-i}) \leq b | v_i] \mu(db_{-i}) > 0 \forall v_i \in [0, \bar{v}] \right\}. \]

Then beliefs for each player \( i \) belong to \( \Delta^+(B_{-i}) \).

Focusing attention to \( \Delta^+(B_{-i}) \) 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 (1.2.2) rule out weakly dominated bids\(^9\). This also yields a natural upper bound on rationalizable bids which will be the starting point of the iterative method described below.

1.2.1.3 Rationalizable bids for a given upper bound on opponents’ bidding functions

Let \( \overline{B} : [0, \bar{v}] \to \mathbb{R}_+ \) 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 \( v \in [0, \bar{v}] \) and each \( j \neq i \). Denote \( \overline{B}_{-i} = \{ \overline{B}, \overline{B}, \ldots, \overline{B} \} \). Accordingly, let
\[ \Delta^+(B_{-i}; \overline{B}_{-i}) = \left\{ \mu \in \Delta^+(B_{-i}): \mu \left( \{ b_{-i}: b_{-i} < \overline{B}_{-i} \} \right) = 1 \right\}. \]

Which bids can be rationalized as best responses to some beliefs in \( \Delta^+(B_{-i}; \overline{B}_{-i}) \)?

---

\(^9\)In the second-price auctions it implies that players bid their private values.
payoff for bid $b^*$) against $\inf_{\mu \in \Delta^+(B_{-i}, \overline{B}_{-i})} \pi^*(v_i; \mu)$ (the worst possible expected payoff for beliefs in $\Delta^+(B_{-i}, \overline{B}_{-i})$). If $v_i - b^* < \inf_{\mu \in \Delta^+(B_{-i}, \overline{B}_{-i})} \pi^*(v_i; \mu)$, then $b^*$ cannot be rationalized as a best-response. From Theorem 6 in B-S it follows that:

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

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

4. If $\overline{B}$ is increasing, then $\inf_{\mu \in \Delta^+(B_{-i}, \overline{B}_{-i})} \pi^*(v_i; \mu) = \pi^*(v_i; \overline{B}_{-i})$, 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 $\overline{B}_{-i}$.

### 1.2.1.4 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 (1.2.1) and (1.2.2), 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 dominated strategies.

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

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

denote the set of bids that can be rationalized as best-responses for bidder $i$ with private value $v_i$ to some beliefs in $\Delta_i$. Next, for a collection of set-valued functions $\{C_j: [0, \overline{v}] \Rightarrow \mathbb{R}_+\}_{j \neq i}$, let denote the set of beliefs allowed by the assumption (1.2.2) with the support in $C_{-i} \equiv \prod_{j \neq i} C_j$ by

$$\Delta^+(C_{-i}) = \{\mu \in \Delta^+(B_{-i}): \mu(\{b_j(v_j) \in C_j(v_j), \forall v_j \in [0, \overline{v}], \forall j \neq i\}) = 1\}.$$  

10I consider interim $k$-rationalizability, so a belief $\mu_i$ can depend on the private value $v_i$. 


Definition 1.2.1 (k-rationalizable and rationalizable bids) For each $i = 1, \ldots, n$ and each $v_i \in [0, \bar{v}]$, let:

\[
\begin{align*}
R_{i,0}(v_i) &= \mathbb{R}_+ , \\
R_{i,k}(v_i) &= \rho_i \left( v_i, \Delta^+ (R_{-i,k-1}) \right), \quad k = 1, 2, \ldots
\end{align*}
\]

1. A bid $b^*$ is $k-$rationalizable for bidder $i$ given $v_i$ if $b^* \in R_{i,k}(v_i)$.

2. A bid $b^*$ is rationalizable for bidder $i$ given $v_i$ if there exists an $n-$tuple $(C_1, \ldots, C_n)$ such that $C_j(v_j) \subseteq \rho_j(v_j, \Delta(C_{-j}))$ for each $j = 1, \ldots, n$, and $b^* \in C_i(v_i)$.

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. As it turns out, that under assumptions (1.2.1), (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 (1.2.1), (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 $B_k(\cdot)$. For every $v_i \in [0, \bar{v}]$, let

\[
\begin{align*}
B_1(v_i) &= v_i , \\
B_{k+1}(v_i) &= v_i - \inf_{\mu \in \Delta^+(\mathbb{R}_{-i,k}(\cdot))} \pi^*(v_i; \mu), \quad k = 1, 2, \ldots
\end{align*}
\]

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

\[
\inf_{\mu \in \Delta^+(\mathbb{R}_{-i,k}(\cdot))} \pi^*(v_i; \mu) = \pi^*(v_i; B_k(\cdot)),
\]

where

\[
\pi^*(v_i; B_k(\cdot)) = \sup_{b \geq 0} \left\{ (v_i - b) \mathbb{P} \left[ B_k(v_j) \leq b \forall j \neq i \right] \right\}
\]
And one can re-express these bounds simply as
\[B_1(v_i) = v_i, \]
\[B_{k+1}(v_i) = v_i - \pi^*(v_i; B_k(\cdot)), \quad k = 1, 2, \ldots. \tag{1.2}\]

**Result 1.2.1 (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.

4. The “minimum shading” function \(S_k(v_i) \equiv v_i - \overline{B}_k(v_i)\) is increasing (non-decreasing for \(k = \infty\)) and convex, with \(S_k(0) = 0\) and \(S_k(v_i) > 0\) for all \(v_i \in (0, \overline{v}]\).

**1.2.2 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 1.2.2, 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\(^{11}\) (see Crawford and Iriberri (2007), Crawford et al. (2013), Kline (2015)).

\(^{11}\)Typically, those models are based on specific assumptions about “anchor beliefs” for a strategically naive “level-0” type or “\(L0\)” and proceed to adjust beliefs also in an iterative way, so that \(L1\) best responds to \(L0\), \(L2\) to \(L1\), and so on. If the anchor beliefs of the \(L0\) type are consistent with Assumption 1.2.2 (any bid greater than or equal to reserve price can win the good with positive probability) and if bidders are assumed to be expected-utility maximizers, then \(k\)-rationalizable bounds are valid for each \(Lk\)-level type.
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)).

- **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 bidding functions arises from heterogeneity in beliefs.

The issue of whether learning would inevitably lead to equilibrium bidding in first-price sealed-bid auctions is unclear. The fact that bidders only observe bids and not the actual realization of other bidders’ values means that they obtain limited feedback from the outcomes of previous auctions, possibly preventing or slowing down the convergence to equilibrium and correct beliefs. Furthermore, being able to identify an upper bound for $k$ that is valid for a population of bidders (the main goal of my paper) would be very valuable in trying to make predictions prior to evaluating outcomes for new auction designs for the same object: if the distribution of values is the same and bidders are able to at least identify the initial set of dominated bids, carrying out the iterative procedure up to the upper bound identified for $k$ would constitute a useful prediction for rationalizable bidding functions and outcomes like expected revenue, etc.

As an illustration of the importance of having a bidding behavior model capable of encompassing the above features, consider Figure 2, which compares nonparametric
estimators of bidding functions against a nonparametric estimator of the risk-neutral, symmetric BNE bidding function in first-price auctions of timber tracts. All functions are estimated by conditioning the U.S. Forest Service advertised appraisal value at its mean. The pattern shown in the figure is consistent with bids above and below the risk-neutral BNE. While this is inconsistent with explanations such as risk-aversion, it is entirely consistent with $k$-rationalizable bidding.

**Figure 1.1.** Comparison of nonparametrically estimated bidding functions, BNE bidding functions and $k$-rationalizable bounds for $k = 2$. 

Note: Estimators correspond to first-price auctions with $n = 2$ participants and appraisal value (advertising price) at its average value in the data. Details of their construction can be found in Section 2.
1.2.3 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 or categories of bidders, where bidders are symmetric within each type or category but possibly asymmetric across types in very general ways. To be precise, suppose bidders of type \( r \) have a distribution of private values \( F_r : [0, \tau] \rightarrow [0, 1] \) (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, but they treat all players of the same type symmetrically. Along these lines, bidders’ beliefs may conjecture upper bounds for bids that are type-specific. In such a setting, beliefs for bidder \( i \) is a collection of \( R \) measures \( \{\mu_{ir}\}_{r=1}^R \). The corresponding collection of upper bounds \( \{\mathcal{B}_{ir,k}(\cdot)\}_{r=1}^R \), 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\(^{12} \): \( s \) (for a small firm) and \( \ell \) (for a large firm). The distributions of private values of different types can be different (\( F_s \) is not necessarily equal \( F_\ell \)). \( F_s \) and \( F_\ell \) are common knowledge\(^{13} \). I do not impose any assumptions on stochastic dominance between \( F_s \) and \( F_\ell \).

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 \( \mu_{ss} \) when her opponent is of type \( ss \) and \( \mu_{sl} \) when her opponent is of type \( \ell \). Suppose \( \mu_{ss} \) and \( \mu_{sl} \) are consistent with \( k_{ss} \) and \( k_{sl} \) steps of elimination of strictly dominated strategies respectively (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

\(^{12}\)This setting is relevant to my empirical illustration with USFS Timber auctions.

\(^{13}\)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, 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.
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. \( k_s \neq k_\ell \). 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 \( \ell \) in the population, there can be three configurations of types of players in an auction with \( n = 2 \) players: \((s, s), (s, \ell)\) and \((\ell, \ell)\). The configurations \((s, s)\) and \((\ell, \ell)\) are just symmetric cases considered in the previous section. I focus on the auction with two different types \((s, \ell)\). As in the symmetric case, the expected payoffs of bidding \( b \), for bidders of type \( s \) and \( \ell \) with private values \( v_s \) and \( v_\ell \) respectively are

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

\[
\pi(b, v_\ell; \mu_{s\ell}) = \int_{B_s} (v_\ell - b) P[b_s \leq b] \mu_{s\ell}(db_s) \quad \text{for bidder } \ell.
\]

Also assume that \( F_s \) and \( F_\ell \) have the same support and (1.2.1) is satisfied for both cdfs. Let \( B_{s\ell,k}(\cdot) \) denote type \( s \)'s \( k \)-rationalizable belief about the upper bound for bidding functions by \( \ell \), and similarly for \( B_{s\ell,k}(\cdot) \). Then,

\[
\begin{aligned}
B_{s\ell,1}(v_\ell) &= v_\ell, \\
B_{s\ell,1}(v_s) &= v_s \\
B_{s\ell,k+1}(v_s) &= v_s - \pi^*(v_s; B_{s\ell,k}(\cdot)), \quad k = 1, 2, \ldots \\
B_{s\ell,k+1}(v_\ell) &= v_\ell - \pi^*(v_\ell; B_{s\ell,k}(\cdot)), \quad k = 1, 2, \ldots
\end{aligned}
\]

where

\[
\pi^*(v_s; B_{s\ell,k}(\cdot)) = \sup_{b \geq 0} \{(v_s - b) P[B_{s\ell,k}(v) \leq b]\}
\]

\[
= \sup_{b \geq 0} \left\{(v_s - b) \int_v 1 \left[ B_{s\ell,k}(v) \leq b \right] dF_\ell(v) \right\},
\]

\[
\pi^*(v_\ell; B_{s\ell,k}(\cdot)) = \sup_{b \geq 0} \{(v_\ell - b) P[B_{s\ell,k}(v_s) \leq b]\}
\]

\[
= \sup_{b \geq 0} \left\{(v_\ell - b) \int_v 1 \left[ B_{s\ell,k}(v) \leq b \right] dF_s(v) \right\}
\]

The assumption that both \( F_s \) and \( F_\ell \) have the same support would prevent the possibility of the \( k \)-rationalizable bounds may have flat parts. All other properties
1.3 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 $k^{14}$.

1.3.1 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 (1.2.1) and (1.2.2), bidding behavior is consistent with $k$—rationalizability only if $b_i \leq B_k(v_i)$ w.p.1. for each bidder $i^{15}$. This implies a first-order stochastic dominance condition:

$$P(B_k(v_i) \leq t) \leq P(b_i \leq t) \quad \forall \ t.$$  

I focus on the setting described in Section 1.2. 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

$$B_k(v) = t$$

has a unique solution in $v$. Denote this solution as $\overline{v}_k(t)$. Then, the above inequality can be expressed as

$$P(v_i \leq \overline{v}_k(t)) \leq P(b_i \leq t) \quad \forall \ t \in [\underline{b}, \overline{b}]^{16} \quad (1.3)$$

I could base a test for $k$—rationalizable bidding in the population on the stochastic dominance condition described in (1.3) if I knew (or could identify):

1. The distribution of values.

2. The distribution of bids

---

14 In my empirical application, I consider an extension to two types in two-bidders auctions.
15 If all bidders have the same beliefs, the statement is "if and only if".
16 The support of bids can depend on $X$ and be the type $[\underline{b}(X), \overline{b}(X)]$.
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.

### 1.3.2 A nonparametric test for $k$-rationalizability with auxiliary data from ascending auctions

The stochastic dominance restriction in (1.3) 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.

A stylized version of an ascending auction is the so-called *button auction* (Milgrom and Weber (1982)) where bidders hold down a button to remain active as the price rises continuously, releasing the button to drop out irreversibly until only one bidder remains. Thus, each losing bidder’s valuation would be learned exactly from her bid. Paarsch (1997) estimated a parametric model of ascending auctions under the button auction framework and IPV. Athey and Haile (2002) showed that in button auctions, transaction prices alone are sufficient to nonparametrically identify the distribution of bidders’ values in an IPV model. Haile and Tamer (2003) analyzed “incomplete” models of ascending auctions in a more general setting relying only on two behavioral assumptions: (i) bidders do not bid more than their valuations, and (ii) bidders do not allow an opponent to win the good at a price they are willing to pay. They showed that in the IPV framework, these assumptions are sufficient to nonparametrically identify bounds for the underlying distribution of valuations. I assume the following about the auctions observed.

**Assumption 1.3.1** 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\[17\] $[\underline{v}, \overline{v}]$.

\[17\text{In all the results that follow, the support can depend on } X \text{ and be of the type } [\underline{v}(X), \overline{v}(X)].\]
Remark 1.3.1

(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 1.3.1 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.

(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 $\xi$ that is independent of the private values conditional on $X$ ($\xi$ 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 (1.2.1) and (1.2.2) 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 1.3.1 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.

\footnote{I do not allow for unobserved heterogeneity at the auction level. However, correlation in bids in first-price auctions is possible in my model if beliefs are correlated, perhaps through a publicly observed signal that affects beliefs but not values.}
Assumption 1.3.2 The transaction price observed in the ascending auctions corresponds to the maximum between the reserve price and the second-highest bidder’s valuation.

Assumption 1.3.2 corresponds exactly 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

Assumption 1.3.3

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 1.3.1. Let $F_{2,P|X}(.|X)$ 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 $(b^j_i)_{i=1}^{n_j}$, the collection of all bids submitted by the $n_j$ bidders. Conditional on $X_j = x$ (with $n_j = n \geq 1$), bids $(b^j_i)_{i=1}^{n_j}$ are assumed to be identically distributed, with cdf $G_1(.|X)$. Specifically,

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

Remark 1.3.2

1. Under the assumption of symmetric bidders, it is enough to observe transaction price (or a randomly drawn bid) for each sealed-bid auction. Conversely, if bids were not observed, one would need to observe all bids to estimate the cdf of bids.

Jump bidding occurs when observed changes in bids are above the minimal allowed increment. The presence of jump bids can be studied empirically by comparing the relative difference between the two highest bids.

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

Bids $(b^j_i)_{i=1}^{n_j}$ 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 $X$. 

---

\(^{19}\)Jump bidding occurs when observed changes in bids are above the minimal allowed increment. The presence of jump bids can be studied empirically by comparing the relative difference between the two highest bids.

\(^{20}\)Under the assumption of symmetric bidders, it is enough to observe transaction price (or a randomly drawn bid) for each auction.

\(^{21}\)Bids $(b^j_i)_{i=1}^{n_j}$ 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 $X$. 

19
other than transaction price are also observed in the ascending auctions data, they can be included in the analysis.

2. The most natural way to interpret the restriction that bids \((b^i_j)_{i=1}^{n_j}\) are identically distributed conditional on \(X_j\) is that bids can be written as \(b^i_j = b(v^i_j; X_j, \mu(v^i_j; X_j, \xi^i_j))\), where the function \(b(\cdot)\) is the same across bidders\(^{23}\), \(v^i_j\) denotes \(i\)'s valuation, \(\xi^i_j\) is unobserved heterogeneity that determines \(i\)'s beliefs and \((v^i_j, \xi^i_j)_{i=1}^{n_j}\) 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 ascending auctions, respectively. Then \(E_{1,X}[g(X)] = \int g(x) dF_{1,X}(x), E_{2,X}[g(X)] = \int g(x) dF_{2,X}(x), \) 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 1.3.2 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 1.3.1 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

\[ \Omega_n\left(F_0(v|x); F_{2,P|X}(v|x)\right) = 0, \quad \forall \, v \in [\underline{v}, \overline{v}], x \in S_{2,X} \]  \hspace{1cm} (1.4)

\(^{22}\)Apart from \(v\), the only source of unobserved heterogeneity is beliefs \(\mu\). \(\xi\) can be of any dimension, even a function.

\(^{23}\)In the asymmetric case with more than one type of players, bidding function \(b\) and belief \(\mu\) can be indexed by the type \(t\), i.e. \(b^i_{j,t} = b_t(v^i_j; X_j, \mu_t(v^i_j; X_j, \xi^i_j))\), where \((v^i_j, \xi^i_j)\) 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\footnote{With a nonbinding reserve price (i.e, $r \leq v_i$), w.p.1}

\[
\overline{B}_1(v_i|x) = v_i, \\
\overline{B}_{k+1}(v_i|x) = v_i - \pi^*(v_i; \overline{B}_k(\cdot|x), x), \quad k = 1, 2, \ldots
\]

where

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

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

The stochastic dominance condition in (1.3) becomes

\[
F_0(\overline{v}_k(t|x)|x) \leq G_1(t|x) \quad \forall \ t \in [b, \overline{b}], x \in S_{2,X} \cap S_{1,X} \tag{1.6}
\]

1.3.2.1 An econometric test

In this section, I describe an econometric procedure to test $k$–rationalizability based on the stochastic dominance condition (1.6). Econometric inference and testing involving inequalities of conditional functionals (moments, distributions, etc.) have been studied in Ghosal et al. (2000), Barrett and Donald (2003), Hall and Yatchew (2005), Lee et al. (2009), Andrews and Shi (2013), Andrews and Shi (2011), Chernozhukov et al. (2013), Lee et al. (2013), Ponomareva (2010), Kim (2009), Menzel (2014), Armstrong (2015) Armstrong (2014), Chetverikov (2017), Lee et al. (2014), and Aradillas-López et al. (2016). 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 (1.6) 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-López 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 (a potentially challenging task when $x$ includes a large collection of covariates). The approach in Aradillas-López 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-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 (1.6). 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 (1.6).

**Assumption 1.3.4** Denote the support of bids as $[b, b]$. 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 $\mathcal{X}$ be a compact subset of $\text{int}(S_{1,X}) \cap \text{int}(S_{2,X})$, and let $\mathcal{B}$ be a compact subset of $\text{int}([b, b])$. Define

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

These sets are assumed to satisfy the following conditions:

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

(ii) $0 < \underline{c} \leq F_0(v|x) \leq \overline{c} < 1$ and $0 < \underline{c} \leq F_{2,p|X}(v|x) \leq \overline{c} < 1$ for all $v \in \mathcal{V}_k$, $x \in \mathcal{X}$.
I rewrite restriction (1.6) in terms of a mean-zero condition. For a given \( t \in [b, \bar{b}] \), \( x \in S_{2,X} \) and \( k \geq 1 \), define

\[ \phi_k(t|x) \equiv F_0(v_k(t|x)|x) - G_1(t|x). \]

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 \( t \in [b, \bar{b}] \); in particular \( \phi_k(t|x) \leq 0 \) for all \((x, t)\) such that \( x \in \mathcal{R} \) and \( t \in \mathcal{B}(x) \).

This is the basis of my test.

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

\[ \Lambda_k(x) \equiv \int_{t \in \mathcal{B}(x)} (\phi_k(t|x))_+ dQ(t|x), \]

\[ T_k \equiv E_{2, X} [\Lambda_k(X) \cdot W_\mathcal{R}(X)] + E_{1, X} [\Lambda_k(X) \cdot W_\mathcal{R}(X)] \tag{1.7} \]

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

A nonparametric estimator for \( T_k \)

I propose a nonparametric test-statistic along the lines of Lee et al. (2013), Lee et al. (2014) and Aradillas-López 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 1.3.5** The vector of auction characteristics \( X \) is partitioned as \( X \equiv (X^c, X^d) \), 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 \( p_{2,X^d|X^c}(x^d|x^c) = P_{2,X}(X^d = x^d|X^c = x^c) \) and
\[ p_{1,X^d|X^c}(x^d|x^c) = P_{1,X}(X^d = x^d|X^c = x^c). \] 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 1.3.9.

### 1.3.2.2 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

\[ \mathcal{H}(X_i - x; h_{x,L_2}) = K \left( \frac{X_i^c - x^c}{h_{x,L_2}} \right) 1 \left[ X_i^d = x^d \right], \]

\[ \widehat{f}_{2,X}(x) = (L_2 h_{x,L_2}^c)^{-1} \sum_{i=1}^{L_2} \mathcal{H}(X_i - x; h_{x,L_2}), \]

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

\[ \widehat{f}_{2,P|X}(p|x) = \frac{\widehat{f}_{2,(p,X)}(p, x)}{\widehat{f}_{2,X}(x)}, \]

\[ \widehat{F}_{2,P|X}(p|x) = \int_{-\infty}^{p} \widehat{f}_{2,P|X}(t|x) dt \]

\[ = \frac{(L_2 h_{x,L_2}^c)^{-1} \sum_{i=1}^{L_2} \left( \int_{-\infty}^{p} \frac{1}{h_{p,L_2}} k \left( \frac{P_i - t}{h_{p,L_2}} \right) dt \right) \mathcal{H}(X_i - x; h_{x,L_2})}{\widehat{f}_{2,X}(x)}. \]
My choice of estimator for $\hat{F}_{P|X}(\cdot|x)$ helps simplify the asymptotic analysis because its argument is itself a nonparametric estimator. Note that using Leibniz rule,

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

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 (1.4) after I replace $F_{P|X}(v|x)$ with $\hat{F}_{P|X}(v|x)$. My estimator $\hat{F}_0(v|x)$ is the solution, in $s$, to the equation $\Omega_n \left( \hat{F}_0(v|x); \hat{F}_{P|X}(v|x) \right) = 0$. Therefore $\hat{F}_0(v|x)$ is defined implicitly by the equation

$$\Omega_n \left( \hat{F}_0(v|x); \hat{F}_{P|X}(v|x) \right) = 0. \quad (1.8)$$

**Estimation of the $k$–rationalizable bounds**

I estimate the bounds with sample of ascending auctions analogs to the construction in (1.5). For a given $v$ and $x$, I have

$$\hat{B}_1(v|x) = v,$$

$$\hat{B}_{k+1}(v|x) = v - \hat{\pi}^*(v; \hat{B}_k(\cdot|x), x), \quad k = 1, 2, \ldots, \quad \text{where} \quad (1.9)$$

$$\hat{\pi}^*(v; \hat{B}_k(\cdot|x), x) = \sup_{b \geq 0} \left\{ (v - b) \left( \hat{F}_0 \left( \hat{\pi}_k(b|x) \right) \right) \right\},$$

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

$$\hat{b}^*_k(v|x) = \arg\max_{b \geq 0} \left\{ (v - b) \left( \hat{F}_0 \left( \hat{\pi}_k(b|x) \right) \right) \right\},$$

$$\hat{\pi}_k^*(v|x) \equiv \hat{\pi}_k(\hat{b}^*_k(v|x)|x),$$

Then,

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

$$
\hat{f}_{1,X}(x) = (L_1 h_{x,L_1}^c)^{-1} \sum_{j=1}^{L_1} \mathcal{H}(X_j - x; h_{x,L_1}),
$$

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

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

$$
\hat{G}_1(b|x) = \int_{-\infty}^{b} \hat{g}_{1,b|x}(t|x) dt
$$

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

1.3.2.4 Estimation of $T_k$

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

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

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

$$
\hat{\Lambda}_{k,1}(x) = \int_{t \in \mathcal{B}(X)} \hat{\phi}_k(t|x) \cdot 1 \left[ \hat{\phi}_k(t|x) \geq -\gamma_{L_1} \right] dQ(t|x),
$$

$$
\hat{\Lambda}_{k,2}(x) = \int_{t \in \mathcal{B}(X)} \hat{\phi}_k(t|x) \cdot 1 \left[ \hat{\phi}_k(t|x) \geq -\gamma_{L_2} \right] dQ(t|x).
$$

The rate-of-convergence restrictions for $\gamma_{L_1}$ and $\gamma_{L_2}$ are described in Assumption 1.3.9, below. For a testing range $\mathcal{X}$, $\mathcal{B}$ and for the weighting functions $Q$ and $W_{\mathcal{X}}$,
I estimate
\[ \hat{T}_k = \frac{1}{L_2} \sum_{i=1}^{L_2} \hat{\Lambda}_{k,2}(X_i)W_{\mathcal{X}}(X_i) + \frac{1}{L_1} \sum_{j=1}^{L_1} \hat{\Lambda}_{k,1}(X_j)W_{\mathcal{X}}(X_j), \] (1.10)

The use of the tuning parameters \( \gamma_{L_1} \) and \( \gamma_{L_2} \) generates asymptotic properties for \( \hat{T}_k \) that automatically adapt to the properties of the so-called contact sets, which for a given \( k \) are defined as
\[ \{(t, x) : x \in \mathcal{X}, t \in \mathcal{B}(x), \phi_k(t|x) = 0\}. \]

The contact sets would correspond to regions of \( \mathcal{X} \) where bidders bid at the boundary of the \( k \)--rationalizable bids.

**Assumption 1.3.6** 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 \( x \in \mathcal{X} \), \( f_0(v|x), f_{2,P|X}(v|x), g_{1,|X}(b|x), f_{2,X}(x) \) 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 \( t \in \mathcal{B}(X) \) (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 \( X \in \mathcal{X}^c \)). I describe the condition next.

**Assumption 1.3.7** There exists \( \underline{\gamma} > 0 \) and \( D < \infty \) such that, for all \( t \in \mathcal{B} \),
\[ P_{2,X} (-c \leq \phi_k(t|X) < 0 | X \in \mathcal{X}^c) \leq D \cdot c \quad \forall \ 0 < c \leq \underline{\gamma}, \]
\[ P_{1,X} (-c \leq \phi_k(t|X) < 0 | X \in \mathcal{X}^c) \leq D \cdot c \quad \forall \ 0 < c \leq \underline{\gamma}. \]

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 $\mathcal{B}$, where

$$
\nu_{L_2}(t) = \frac{1}{\sqrt{L_2}} \sum_{i=1}^{L_2} \left( 1[\{-2\gamma_{L_2} \leq \phi_k(t|X_i) < 0\} W_{\mathcal{X}}(X_i) - E_{2,X} [1 \{-2\gamma_{L_2} \leq \phi_k(t|X_i) < 0\} W_{\mathcal{X}}(X_i)] \right),
$$

$$
\nu_{L_1}(t) = \frac{1}{\sqrt{L_1}} \sum_{j=1}^{L_1} \left( 1[\{-2\gamma_{L_1} \leq \phi_k(t|X_j) < 0\} W_{\mathcal{X}}(X_j) - E_{1,X} [1 \{-2\gamma_{L_1} \leq \phi_k(t|X_j) < 0\} W_{\mathcal{X}}(X_j)] \right)
$$

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 $\mathcal{C}(X_0) = \{X_0 \cap C: C \in \mathcal{C}\}$. 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 1.3.8** There exists a $\bar{b} > 0$ such that

$$
\{ \{x \in \mathcal{X}: -\gamma \leq \phi_k(t|x) < 0\} : t \in \mathcal{B}, 0 < \gamma \leq \bar{b} \}
$$

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 1.3.9**

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 3.3.2), meaning that it satisfies

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

$k(\cdot)$ 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,

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

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

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

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

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

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

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

Note that I require that the auxiliary bandwidth $\gamma_{L_q}$ 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 $\gamma_{L_q}$ go to zero faster than $L_q^{-1/2}$. Taken together, my assumptions require that $\gamma_{L_q}$ converge to zero faster than $h_{L_q}$ (the conditions described imply that $(\gamma_{L_q}/h_{x,L_q}^c) \to 0$). I also assume essentially that the sample sizes $L_1$ and $L_2$ grow at a proportional rate, meaning that I rule out $L_2/L_1 \to 0$ and $L_1/L_2 \to 0$. 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 $\hat{G}_1$ and $\hat{F}_{2,P|X}$. The main result for $\hat{T}_k$ is the following.

**Theorem 1.3.1** Let $L \equiv L_1 + L_2$. If Assumptions 1.3.1-1.3.9 are satisfied, then for some $\Delta > \frac{1}{2}$,

$$
\hat{T}_k = T_k + \frac{1}{L_1} \sum_{j=1}^{L_1} \psi_{1,k}(b_j; X_j; h_{b,L_1}, h_{x,L_1}) + \frac{1}{L_2} \sum_{i=1}^{L_2} \psi_{2,k}(P_i; X_i; h_{p,L_2}, h_{x,L_2}) + o_p \left( L^{-\Delta} \right).
$$
where \( \psi_{1,k} \) and \( \psi_{2,k} \) are two influence functions satisfying the following conditions:

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

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

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

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

\[
\lim_{L \to \infty} \sigma_{1,k,L_1}^2 = \sigma_{1,k}^2 > 0 \quad \text{and} \quad \lim_{L \to \infty} \sigma_{2,k,L_2}^2 = \sigma_{2,k}^2 > 0.
\]

**Proof:** A step-by-step proof is included in Econometric Appendix A (the exact expressions for \( \psi_{1,k}(b, X_j; h_{b,L_1}, h_{x,L_1}) \) 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, \( \hat{\phi}_k(t|x) \) has a linear representation of the form

\[
\hat{\phi}_k(t|x) - \phi_k(t|x) = \frac{1}{L_2} \sum_{t=1}^{L_2} \varphi_{0,k}^f (P_i, X_i, t, x; h_{p,L_2}, h_{x,L_2}) - \frac{1}{L_1} \sum_{j=1}^{L_1} \varphi_{1,k}^g (b_j, X_j, t, x; h_{b,L_1}, h_{x,L_1}) + \xi_{L_2}^f (t, x) - \xi_{L_1}^g (t, x), \quad \text{where:}
\]

\[
\begin{align*}
E_{2,(P,X)} [\varphi_{0,k}^f (P_i, X_i, t, x; h_{p,L_2}, h_{x,L_2})] &= 0, \quad \text{for all } t \in \mathcal{B}, x \in \mathcal{X}. \\
E_{1,(b,X)} [\varphi_{1,k}^g (b_j, X_j, t, x; h_{b,L_1}, h_{x,L_1})] &= 0
\end{align*}
\]

\[
\begin{align*}
\sup_{t \in \mathcal{B}} \sup_{x \in \mathcal{X}} |\xi_{L_2}^f (t, x)| &= O_p \left( L_2^{-1/2-\epsilon} \right), \\
\sup_{t \in \mathcal{B}} \sup_{x \in \mathcal{X}} |\xi_{L_1}^g (t, x)| &= O_p \left( L_1^{-1/2-\epsilon} \right).
\end{align*}
\]

The expression for \( \varphi_{0,k}^f \) is derived inductively, starting with \( k = 2 \). From here, the
next step is to show that

\[
\frac{1}{L_1} \sum_{i=1}^{L_1} \hat{\Lambda}_{k,1}(X_i) W_{\mathcal{X}}(X_i) = E_{1,X} [\Lambda_k(X_i) W_{\mathcal{X}}(X_i)] \\
+ \frac{1}{L_1} \sum_{i=1}^{L_1} (\Lambda_k(X_i) W_{\mathcal{X}}(X_i) - E_{1,X} [\Lambda_k(X_i) W_{\mathcal{X}}(X_i)]) \\
+ \frac{1}{L_1 L_2 h_{x,L_2}} \sum_{i=1}^{L_1} \sum_{\ell=1}^{L_2} \int_{t \in \mathcal{B}(X_i)} \varphi^{F_0,k} (P_\ell, X_\ell, t, X_i; h_{p,L_2}, h_{x,L_2}) 1 [\phi_k(t|X_i) \geq 0] dQ(t|X_i) \cdot W_{\mathcal{X}}(X_i) \\
- \frac{1}{L_1^2 h_{x,L_1}^2} \sum_{i=1}^{L_1} \sum_{m=1}^{L_1} \int_{t \in \mathcal{B}(X_i)} \varphi^{G_1} (b_m, X_m, t, X_i; h_{b,L_1}, h_{x,L_1}) 1 [\phi_k(t|X_i) \geq 0] dQ(t|X_i) W_{\mathcal{X}}(X_i) + \hat{\omega}_{k,1},
\]

where \( \hat{\omega}_{k,1} = O_p(L^{-1/2-\epsilon}) \) and \( \hat{\omega}_{k,2} = O_p(L^{-1/2-\epsilon}) \) 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 \( \frac{1}{L_1} \sum_{i=1}^{L_1} \hat{\Lambda}_{k,1}(X_i) W_{\mathcal{X}}(X_i) = O_p(L^{-1/2-\epsilon}) \) and \( \frac{1}{L_2} \sum_{i=1}^{L_2} \hat{\Lambda}_{k,2}(X_i) W_{\mathcal{X}}(X_i) = O_p(L^{-1/2-\epsilon}) \) 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.

**1.3.2.5 A test based on Theorem 1.3.1**

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 (1.6).
The result in Theorem 1.3.1 guides the construction of a test of $k$–rationalizability:

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

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

and therefore $\sqrt{L} \cdot \hat{T}_k \longrightarrow +\infty$ 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 $\mathcal{X} \times \mathcal{B}(x)$, then

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

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

   $$\frac{\hat{T}_k}{\sigma_{k,L_1,L_2}} \overset{d}{\longrightarrow} N(0,1),$$

   where

   $$\sigma^2_{k,L_1,L_2} = \frac{\sigma^2_{1k,L_1}}{L_1} + \frac{\sigma^2_{2k,L_2}}{L_2}.$$  \hfill (1.11)

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 $\sigma^2_{k,L_1,L_2}$. 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 $\lambda_L \rightarrow 0$ 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

$$\bar{T}_k = \frac{\hat{T}_k}{\sqrt{\max \left\{ \sigma^2_{k,L_1,L_2}, \frac{\lambda_k}{L} \right\}}}$$

Then,
1. If bids violate the $k$–rationalizable bounds,

$$
\bar{t}_k = \frac{1}{\sqrt{\max \left\{ \left( \left( \frac{L}{L_1} \right) \sigma_{1_k,L_1}^2 + \left( \frac{L}{L_2} \right) \sigma_{2_k,L_2}^2 \right), \lambda_L \right\}}} \times \sqrt{L} \cdot \hat{T}_k \xrightarrow{\beta + \infty \text{ w.p.1.}} \frac{1}{\sqrt{(\pi^2)\sigma_{1_k}^2 + (\frac{1}{d_2})\sigma_{2_k}^2}}
$$

where $\sigma_{1_k}^2 > 0$ and $\sigma_{2_k}^2$.

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 $\sqrt{L} \cdot \hat{T}_k \xrightarrow{d} N \left( 0, \frac{1}{d_1} \sigma_{1_k}^2 + \frac{1}{d_2} \sigma_{2_k}^2 \right)$, and

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

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

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

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

$$
\hat{t}_k = \frac{\hat{T}_k}{\sqrt{\max \left\{ \hat{\sigma}_{k,L_1,L_2}^2, \frac{\lambda_k}{L} \right\}}}.
$$

(1.12)
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 \( \alpha \), let \( \Phi(z_{1-\alpha}) = 1 - \alpha \), where \( \Phi(\cdot) \) is the Standard Normal distribution. In view of the asymptotic properties of \( \hat{t}_k \), the rejection rule is:

Reject \( H_0 \) if and only if \( \hat{t}_k > z_{1-\alpha} \).

From the asymptotic properties of \( \hat{t}_k \), 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 1.3.1 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)).

1.4 Extensions

1.4.1 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 \( F_{0,\ell} \) 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
\( E_0(\cdot) \) is a lower bound for \( F_0(\cdot) \), so that \( E_0(\cdot|x) \leq F_0(\cdot|x) \) for a.e. \( x \). Suppose \( F_0 \) is not identified but \( E_0(\cdot) \) is. This will be enough to identify an upper bound for \( B_k \), the \( k \)-rationalizable bounds, which I denote as \( \overline{B}_k(\cdot|x) \).

\( k = 1 \): The upper bound \( \overline{B}_1(\cdot|x) \) does not depend on the distribution of private values, and remains as:
\[
\overline{B}_1(v|x) = v
\]

\( k = 2 \): If \( F_0(\cdot|x) \) were identified, the sharp upper bound for 2-rationalizable bids would be given by
\[
\overline{B}_2(\cdot|x) = v - \max_{b \geq 0} (v - b) F_0^{n-1}(b|x),
\]

When \( F_0(\cdot|x) \) is not identified but a lower bound \( F_0(\cdot|x) \) is, the probability of winning the good is bounded below by \( E_0^{n-1}(\cdot|x) \) and corresponds to the worst case scenario for the bidder. From here, an upper bound for 2-rationalizable bids is
\[
\overline{B}_2(v|x) = v - \max_{b \geq 0} (v - b) E_0^{n-1}(b|x)
\]

\( k \geq 3 \): By the same reasoning, an upper bound for any finite \( k \) can be found by induction using the following formula:
\[
\overline{B}_k(v|x) = v - \max_{b \geq 0} (v - b) E_0^{n-1}(\overline{B}_{k-1}^{-1}(b|x))
\]

If \( F_0 \) is a cdf for some distribution and pdf \( f_0 \) exists and positive, the upper bounds \( \overline{B}_k(\cdot|x) \) remain strictly increasing functions that monotonically converge to some limit. Denote the inverse of \( \overline{B}_k(\cdot|x) \) by \( \overline{v}_k(\cdot|x) \). Then the stochastic dominance condition in (1.6) holds only if,
\[
F_0(\overline{v}_k(t|x)|x) \leq G_1(t|x) \quad \forall \ t \in [b, \overline{b}], x \in S_{2,X}
\]

Given my data assumptions, a nonparametric lower bound for \( F_0(\cdot) \) can be obtained using the results from Haile and Tamer (2003). Replace Assumption 1.3.2 with the following restriction.

**Assumption 1.4.1** In ascending auctions, bidders do not allow an opponent to win at a price they are willing to beat.
This corresponds to Assumption 2 in the “incomplete” model of ascending auctions\textsuperscript{25} in Haile and Tamer (2003). It is consistent with Assumption 1.3.2 but it allows for other kinds of bidding behavior, including 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. Take any ascending auction \(i\) with \(n\) bidders, and rank their valuations as order-statistics

\[V_{1:n} \leq V_{2:n} \leq \cdots \leq V_{n-1:n} \leq V_{n:n}\]

Let \(\Delta\) be the minimum required bid increment (treated as observable).

By Assumption 1.4.1, it must be that

\[V_{n-1:n} \leq P_i + \Delta\]

where, as before, \(P_i\) denotes the transaction price in the auction. Since this relationship holds w.p.1, it must be that

\[F_0(v \mid x) \geq \Gamma_n(F_{2,P_i}X(v + \Delta \mid x)) \equiv E_0(v \mid x)\]

and the stochastic dominance condition in (1.6) implies that \textit{bidders are \(k\) rationalizable in first-price auctions only if}

\[E_0(\bar{\nu}_k(t \mid x) \mid x) \leq G_1(t \mid x) \quad \forall \ t \in [\bar{b}, \bar{b}], x \in S_{2,X} \cap S_{1,X}
\]

In this case the test would be based on

\[\phi_k(t \mid x) = E_0(\bar{\nu}_k(t \mid x) \mid x) - G_1(t \mid x)\]

instead of \(\phi_k(t \mid x) = F_0(v_k(t \mid x) \mid x) - G_1(t \mid x)\). The test would be constructed just like before, using the same nonparametric estimators for \(G_1\) and \(F_{2,P_i}X\). Assuming that

\textsuperscript{25}Haile and Tamer (2003) also assume that bidders do not bid more than they are willing to pay, a behavioral restriction already present in my model given the maintained assumption on beliefs: positive bids can win the good with positive probability.
all the conditions imposed for $\phi_k(t|x)$ are satisfied by $\phi_k(t|x)$, the test would have the same asymptotic properties described above. It would have less power because $\phi_k(t|x)$ only detects larger violations of $k$-rationalizability, but the results would be robust to a wider range of behavior in the population of ascending auctions.

1.4.2 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 or second-price auctions. In other words, for any given $k$, any distribution of bids can be rationalizable 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 $F_0 : [0, 1] \rightarrow [0, 1]$.

**Result 1.4.1** 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.

Note first, that if players do not use weakly dominated bids and the support of private values is $[0, 1]$ then the distribution of bids $G(t)$ can not have support $[0, 1]$.\textsuperscript{26} Thus, for a fixed distribution $G(t)$, its support is $[0, 1 - \varepsilon]$ for some fixed $\varepsilon > 0$.\textsuperscript{27}

Next, I show the result (1.4.1) by induction.

For $k = 1$ the upper bound for rationalizable bids is the 45-degree line, which means that the distribution of bids $G(t)$ can be rationalized by the distribution of private values $F_0(t) = G(t)$ for all $t \in [0, 1 - 2\varepsilon]$ and $F_0(t) < G(t)$ for $t \in (1 - 2\varepsilon, 1]$ with $F_0(1) = 1$. Thus, the main stochastic dominance inequality (1.6) is satisfied for any $t \in [0, 1]$.

\textsuperscript{26}If the support for bids is $[0, 1]$, this would imply that bidders with $v = 1$ bid 1. However, this is weakly dominated by a bid $1 - \varepsilon$ for some $\varepsilon > 0$.

\textsuperscript{27}0 is always in the support of $G(t)$, because 0 is the best response if and only if the private values is 0.
Next consider the case $k = 2$. Since there are no assumptions on the distribution of private values $F_0$, let $F_0(t) = t^\theta$, where $\theta > 0$.\footnote{\(F_0(t) = t^\theta\) for $t \in [0, 1]$ is a special case of the Beta-distribution.} Note that $f_0(t) = \theta t^{\theta - 1}$, which is positive.

In the most pessimistic scenario, bidders play best-response to beliefs that opponents are bidding on the upper bound for the previous step (the 45-degree line). Thus, each bidder solves the following maximization problem:

$$\max_{b \geq 0} (v - b) F_{N-1}^0(b) = \max_{b \geq 0} (v - b) b^{\theta(N-1)}.$$  

The optimal bidding strategy in this case is $b^*(v) \equiv \frac{\theta(N-1)}{1 + \theta(N-1)} v$, so that $b^* \in \left[0, \frac{\theta(N-1)}{1 + \theta(N-1)} \right]$. When $\theta$ increases, the optimal bidding strategy $b^*(v)$ converges to the 45-degree line. Since $b^*(v)$ is a best-response to some admissible beliefs for $k = 2$–rationalizability, the upper bound for 2–rationalizable bids is always greater or equal to this bidding function, i.e. $\overline{B}_2(v) \geq b^*(v)$ for all $v \in [0, 1]$. This means that the distribution of bids that results from bidding on the upper bound $\overline{B}_2(v)$ first-order stochastically dominates the distribution of bids that results from bidding according to the strategy $b^*(v)$, i.e.

$$F_0(\overline{B}_2^{-1}(t)) \leq F_0(b^{*-1}(t)), \quad \forall t.$$  

Next, I show that there exists an $\theta > 0$ large enough, such that $F_0(b^{*-1}(t)) \leq G(t)$ over the support of bids $t \in [0, 1 - \varepsilon]$. And therefore, the first-order stochastic dominance inequality (1.6) is satisfied over the entire support of bids $[0, 1 - \varepsilon]$.

Plugging $b^{*-1}(t)$ into $F_0(t)$:

$$F_0(b^{*-1}(t)) = t^\theta \left(1 + \frac{1}{\theta(N-1)}\right)^{\theta(N-1)}.$$  

For large enough values of $\theta$ the following three conditions are satisfied:

1. The highest best response is larger than the highest bid in the support of $G(t)$:

$$1 - \varepsilon < \frac{\theta(N-1)}{1 + \theta(N-1)}. \quad (1.15)$$
2. Since \(g(0) > 0\) and it is continuous, there exists a positive \(\Delta > 0\) such that

\[
\inf_{0 \leq t \leq \Delta} g(t) \geq g > 0.
\]

There exists large enough \(\theta\) such that:

\[
\frac{dF_0(b^*-1(t))}{dt} = \theta t^{\theta-1} \left( \left( 1 + \frac{1}{\theta(N - 1)} \right)^{\theta(N-1)} \right)^{\frac{1}{N-1}} < g, \quad \forall t \in [0, 1 - \varepsilon].
\]

Since the left hand side of the inequality increases with \(t\), it is sufficient that

\[
\theta (1 - \varepsilon)^{\theta-1} \left( \left( 1 + \frac{1}{\theta(N - 1)} \right)^{\theta(N-1)} \right)^{\frac{1}{N-1}} < g.
\]

(1.16)

3. Denote \(\delta \equiv G(\Delta) > 0\). The distribution that results from bidding according to the best response \(b^*\) is smaller than \(\delta\) at the upper bound of the support of bids, \(1 - \varepsilon\):

\[
F_0(b^*-1(1 - \varepsilon)) = (1 - \varepsilon)^{\theta} \left( \left( 1 + \frac{1}{\theta(N - 1)} \right)^{\theta(N-1)} \right)^{\frac{1}{N-1}} < \delta.
\]

(1.17)

The conditions described above are illustrated in Figure 1.2. Since \(G(t)\) and \(F_0(b^*-1(1 - \varepsilon))\) are strictly increasing, inequalities (1.15), (1.16) and (1.17) guarantee that \(F_0(b^*-1(t)) \leq G(t)\) for all \(t \in [0, 1 - \varepsilon]\). Thus, first-order stochastic dominance inequality is satisfied for all bids in the support.

The proof can proceed by induction for \(k \geq 3\) by using the same argument, since, as I showed before, the bound for \(k = 2\) can become arbitrary close to the 45-degree line.

Result 1.4.1 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 \(F_\theta\) have the same support \([0, 1]\) (for simplicity). In the IPV model, as I showed in Section 1.4.1, we can focus on a lower-envelope for this family, \(E_\theta(v) \equiv \inf_{F \in \mathcal{F}} F(v)\). This lower envelope can be used to construct
Figure 1.2. Graphical illustration of conditions (1.15), (1.16) and (1.17) in the proof of Result 1.4.1

an upper envelope for the bound $\overline{B}_k(v)$ (denoted as $\overline{B}_k(v)$). The features of $\mathcal{F}$ will determine whether $\overline{B}_k(v)$ will be nontrivial (i.e., whether it will be bounded away from the 45-degree line) and also whether this bound is sharp.

If the family $\mathcal{F}$ satisfies the following requirements:

1. For any $F \in \mathcal{F}$, there exists $f = F'$.

2. There exist two strictly positive and bounded constants $C_1$ and $C_2$ such that $\forall F \in \mathcal{F}, C_1 \leq f \leq C_2$ on its support.

3. Any two distributions in $\mathcal{F}$ are mutually absolutely continuous.

Then $F_0$ is a distribution function with density $f_0$'. In addition, $C_1 \leq f_0 \leq C_2$ on its support $[0, 1]$ (See Lemma 1 in Aryal et al. (2016)).

**Remark 1.4.1** If the lower envelope $F_0$ is a member of the parametric family $\mathcal{F}$, then the upper envelope $\overline{B}_k$ 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 $\sup_{v \in [0,1]} \left| \overline{B}_k(v) - \overline{B}_k^F(v) \right| < \epsilon$. Sharpness can result even if $F_0 \notin \mathcal{F}$. 

40
The following three examples illustrate the importance of the choice of the parametric family $\mathcal{F}$

1. If $\mathcal{F} = \{F_\theta(t) = t^\theta, \text{ s.t. } \theta > 0\}$, the lower envelope is $\underline{F}_0(t) = 0$ for $t \in [0, 1)$ and $\underline{F}_0(1) = 1$ (this is a case where condition 2 described above fails, since $f_\theta(t) = \theta t^{\theta-1}$ can be arbitrarily large if any $\theta > 0$ belongs in the family). As I showed in the proof of Theorem 1.4.1, $k$–rationalizability can not be rejected for any finite $k$ for this family.

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

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

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 $\overline{B}_k$ are sharp and informative. This example is illustrated in Figure 1.3.

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 (the conditions for sharpness would fail for $\nu > 0.5$).

If $\overline{B}_k$ is not a sharp bound, or if we want to conduct inference both on $\theta$ and $k$, we can use the conditional moment inequalities implied by $k$–rationalizability and proceed as described in footnote 35 (below).

1.4.3 $k$-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 1.4.2 (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 $E_0(t) \equiv \inf_{f \in \mathcal{F}} F(t).$ 29 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 $E_0$ is identified but perhaps $F_0$ is not.

**Assumption 1.4.3** Bidders solve ambiguity related to the distribution of private values by best-responding to the most pessimistic belief given by the lower envelope $E_0$

If assumptions 1.4.2 and 1.4.3 are satisfied then bidder $i$ solves the following maximization problem given beliefs $\mu_i$ (see p.6 in Aryal et al. (2016)):

$$\pi^*(v_i; \mu_i) = \max_{b \geq 0} (v_i - b) P_{E_0}(b_{-i} < b | \mu_i)$$

If the lower envelope $E_0$ is a cdf and there exists $f_0 = F_0' > 0$ in the support of private values, then the upper bound of $k$–rationalizable bids $B_k(\cdot)$ for any given $k$ can be found by induction:

$$B_1(v) = v,$$

$$B_k(v) = v - \max_{b \geq 0} (v - b) E_0^{k-1}(B_{k-1}(b)). \quad (1.18)$$

**Remark 1.4.2** In contrast to the semi-parametric approach in section 1.4.2, with ambiguity under Assumption 1.4.3, $B_k(\cdot)$ is always a sharp upper bound for $k$–rationalizable bids.

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

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

If $F_0$ is not identified30 , the only testable implication is:

---

29One 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 1.4.2, it is enough to have common lower envelope $E_0$. The estimator of $F_0$ was described in Section 1.4.1

30In contrast to Aryal et al. (2016) the distribution of private values $F_0$ and the lower envelope of the prior set $E_0$ can not be identified from the sample of first-price auctions since I do not assume equilibrium behavior.
\[ F_0(\overline{B}_k^{-1}(t)) \leq G(t), \quad \forall t \in [b, \overline{b}]. \]

**Assumption 1.4.3’** 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 1.4.2, 1.4.3 and 1.4.3’ are satisfied then for any fixed \( k \) players bid according to the bidding function:

\[ \tilde{b}_k^*(v) = \arg\max_{b \geq 0}(v - b)F_0^{n-1}(\overline{B}_{k-1}(b)), \]

where \( \overline{B}_k \) is defined by the iterative procedure (1.18).

In this case the model implies the following equality:

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

If only the lower envelope \( F_0 \) is identified, I can only test the inequality:

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

**Assumption 1.4.3”** 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 1.4.2, 1.4.3 and 1.4.3” are satisfied then for any fixed \( k \) players bid according the bidding function defined iteratively:

\[ \tilde{b}_k^*(v) = v, \]

\[ \tilde{b}_k^*(v) = \arg\max_{b \geq 0}(v - b)F_0^{n-1}(\tilde{b}_{k-1}^{*^{-1}}(b)). \]

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

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

The implications in all three cases can be tested similarly to (1.6).

---

31 If an upper bound \( \overline{F}_0 \) is also identified, as it is in Haile and Tamer (2003). Then another testable implication is \( G(t) \leq \overline{F}_0(\tilde{b}_k^{*^{-1}}(t)), \forall t \in [b, \overline{b}] \). The test for the inequalities \( F_0(\tilde{b}_k^{*^{-1}}(t)) \leq G(t) \leq \overline{F}_0(\tilde{b}_k^{*^{-1}}(t)), \forall t \in [b, \overline{b}] \) can be constructed similar to the one in Section 2.5.3.

32 An upper bound can also be used as described in Footnote 31.
1.4.4 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. As a result, existing nonparametric identification results in first-price auctions cannot be applied here.

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.

(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 $F_s^0(\cdot)$ 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 $F_s^0|X(\cdot|X)$ denote the joint cdf of signals conditional on $X$.

---

33Nonparametric 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.

34Result 1.4.1 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.
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 $E[v(s_i, s_{-i})|s_i, b_{-i} \leq b] \leq E[v(s_i, s_{-i})|s_i]$ (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 (1.5), replacing bidders’ unobserved values with their conditional expectations given the signals. The iterative construction in (1.19) now becomes

$$
\begin{align*}
\bar{B}_1(s_i|X) &= E[v(s_i, s_{-i})|s_i, X], \\
\bar{B}_{k+1}(s_i|X) &= E[v(s_i, s_{-i})|s_i, X] - \pi^*(s_i; \bar{B}_k(\cdot|X), X), \quad k = 1, 2, \ldots, \\
\pi^*(s_i; \bar{B}_k(\cdot|X), X) &= \\
&\sup_{b \geq 0} \left\{ \int_{\max_{j \neq i} \bar{B}_k(s_j|X) \leq b} (E[v(s_i, s_{-i})] - b) \, dF_{s_{-i}|s_i, X}(s_{-i}|s_i, X) \right\} \cdot P\left(\max_{j \neq i} \bar{B}_k(s_j|X) \leq b\bigg| s_i, X\right).
\end{align*}
$$

(1.19)

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 $\Theta$ that satisfies the symmetry and affiliation requirements described above. For a given $\theta$, a parametric expression for $\bar{B}_k(s_i|X; \theta)$ for the corresponding bounds is given by (1.19). 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 $\bar{s}_k(\cdot|x, \theta)$ denote the inverse function of $\bar{B}_k(\cdot|x, \theta)$, and for a given $t \in \text{Supp}(S)$ define

$$
\phi^*_k(t|x, \theta) = F_0^s(\bar{s}_k(t|x, \theta)|x, \theta_2) - G_1(t|x),
$$

where, as before, $G_1(\cdot|X)$ represents the (nonparametrically specified) cdf of bids conditional on $X$. $k-$rationalizability requires that $\phi^*_k(t|x, \theta) \leq 0$ for all $(t, x)$. Similar to (1.7), I can carry out inference based on these stochastic dominance conditions implied by the model.

Denote $\phi^*_k(t|x) \equiv \inf_{\theta \in \Theta} F_0^s(\bar{s}_k(t|x, \theta)|x, \theta_2) - G_1(t|x)$. The following inequality:
\[ \phi_k^s(t|x) \leq 0, \quad \forall (t,x) \quad (1.20) \]

is a necessary condition for \( k\)-rationalizability, i.e. if there exists such \( \tilde{\theta} \in \Theta \) that \( \phi_k^s(t|x,\tilde{\theta}) \leq 0 \) then the inequality (1.20) is satisfied. It can be tested similarly to (1.6)\(^{35}\).

### 1.4.5 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 \( B_1(v) = v \). 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 \( B_1(v) \). 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:

\(^{35}\)For a fixed \( k \), the set of parameters \( \Theta_1 \) 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 \( k \).
\[ \pi^*(v) = \max_{b \geq 0} u(v - b)F_0(b) = \max_{b \geq 0} (v - b)^\alpha F_0(b). \]

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

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

Thus, the upper bound for \( k = 2 \) is:

\[ \overline{B}_2(v) = v - \alpha \frac{F_0^{\frac{1}{\alpha}+1}(b^*(v))}{f_0(b^*(v))}. \quad (1.21) \]

When \( \alpha \) 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\(^{36} \). So, if the parameter of relative risk-aversion \( r = 1 - \alpha \) is not bounded away 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 (1.21):

\[ \overline{B}_2(v) = v - \frac{\alpha \theta \pi}{(\theta + \alpha)^{\frac{1}{\alpha}}} v^\frac{\alpha + \theta}{\alpha}. \]

When \( \alpha \) decreases to zero and \( \theta \) increases to infinity, the upper bound converges to 45-degree line. Similar to the result 1.4.1, 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 1.4.

\(^{36}\)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.
Figure 1.3. Truncated exponential family example. Distribution, lower envelope and $k-$rationalizable bounds (for $k = 6$).
Figure 1.4. $k$–rationalizable bounds with risk-aversion (for $k = 6$). An illustration with CRRA utility function and $U[0,1]$ valuations.
1.4.6 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.

1.4.7 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 1.4.1, 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 $\hat{F}_{0,2}(\cdot|x)$ would still control for size (but may be conservative) if $F_{0,2}(\cdot|x) \leq F_{0,1}(\cdot|x)$ for a.e $x \in \mathcal{X}$. 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 \( X \) and \( X' \) such that

\[
x \in X, \ x' \in X' \implies F_{0,2}(\cdot|x) \leq F_{0,1}(\cdot|x').
\] (1.22)

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 (1.13) would yield now

\[
F_{0,2}(v_{k,2}(t|x'|x')) \leq F_{0,1}(v_{k,1}(t|x)|x) \leq G_1(t|x),
\]

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

\[
\phi_k(t|x, x') \cdot 1[x \in X, x' \in X'] \leq 0, \ \forall (x, x', t).
\]

A condition like (1.22) 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.

### 1.4.7.1 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 1.3.1, 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,

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

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
The test for $k$-rationalizability is still valid but can be conservative.

### 1.5 Concluding remarks

In this paper I analyze testable implications of strategically sophisticated bidding in sealed-bid 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 dominated strategies. $k$-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.
Chapter 2 | Testing $k$–rationalizability in USFS timber auctions

2.1 Introduction

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 1.3.2. 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). I first provide 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). 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”($\ell$) according to the number of workers, with the threshold being 150 workers. In the data, the proportion of observations with two small is 44%, with two large firms is 27% and one small and one large firm is 28%. 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. To make things tractable, I focus on the population of 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 $((\ell, \ell), (\ell, s), (s, s))$.

My results show that values of $k$ as low as 2 can be rejected when two bidders of different type face each other, while bidding behavior is consistent with larger values of $k$ when two bidders of the same type face each other. The results indicate that large bidders underestimate the degree of rationality (as measured by $k$) of small bidders. 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 2.5 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.

2.2 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. During the period of my sample I can also compute measures of bidder inventory as well as the number of previous auctions where each bidder participated (and the ones they won). The data also contains 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 the tract to take into account the diversity of the species), road construction costs, logging costs, manufacturing costs, total costs. For every auction I also observe the USFS advertised value (reserve price), the date of the auction and the contract length. In the case of first-price auctions I observe all bids submitted, and in the case of 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 1.3.2 describes. Summary statistics for all auctions and for the subpopulation of ascending and sealed-bid auctions are given in Table (2.1)
Table 2.1. Summary statistics

<table>
<thead>
<tr>
<th></th>
<th>All auctions 3387</th>
<th>Ascending auctions 2304</th>
<th>Sealed auctions 881</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. dev.</td>
<td>Mean</td>
</tr>
<tr>
<td>Transaction price ($/mbf)</td>
<td>220.29</td>
<td>257.91</td>
<td>207.38</td>
</tr>
<tr>
<td>Reserve price($/mbf)</td>
<td>184.74</td>
<td>182.65</td>
<td>173.72</td>
</tr>
<tr>
<td>Number of bidders</td>
<td>2.09</td>
<td>1.84</td>
<td>2.12</td>
</tr>
<tr>
<td>Volume (mbf)</td>
<td>2093.26</td>
<td>2855.22</td>
<td>2502.57</td>
</tr>
<tr>
<td>Herfindahl index</td>
<td>0.75</td>
<td>0.24</td>
<td>0.74</td>
</tr>
<tr>
<td>Road construction ($/mbf)</td>
<td>9.62</td>
<td>29.86</td>
<td>12.02</td>
</tr>
<tr>
<td>Logging costs ($/mbf)</td>
<td>323.41</td>
<td>125.27</td>
<td>329.46</td>
</tr>
<tr>
<td>Manufacturing costs ($/mbf)</td>
<td>290.55</td>
<td>129.67</td>
<td>287.31</td>
</tr>
<tr>
<td>Density (mbf/acre)</td>
<td>12.01</td>
<td>27.92</td>
<td>11.47</td>
</tr>
<tr>
<td>Contract duration (days)</td>
<td>653.46</td>
<td>604.37</td>
<td>704.25</td>
</tr>
</tbody>
</table>

My test allows the observable characteristics $X$ to have different distributions across both 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 (0 for Oregon state and 1 for Washington state) and include dummies for years and quarters of the sale. The results of the logit regression are represented in the table (2.2).

<table>
<thead>
<tr>
<th></th>
<th>Logit Coefficient</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reserve price ($/mbf)</td>
<td>0.0027***</td>
<td>(0.0004)</td>
</tr>
<tr>
<td>Volume mbf</td>
<td>-0.0003***</td>
<td>(0.0000)</td>
</tr>
<tr>
<td>Road construction ($/mbf)</td>
<td>-0.0019</td>
<td>(0.0011)</td>
</tr>
<tr>
<td>Logging costs ($/mbf)</td>
<td>-0.0001</td>
<td>(0.0002)</td>
</tr>
<tr>
<td>Manufacturing costs ($/mbf)</td>
<td>-0.0004</td>
<td>(0.0003)</td>
</tr>
<tr>
<td>Herfindahl index</td>
<td>0.8460***</td>
<td>(0.2600)</td>
</tr>
<tr>
<td>State</td>
<td>1.9860***</td>
<td>(0.1190)</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

From Table 2.2, 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 auction format is entirely consistent with the assumptions of my test, which only require that, for a given (fixed) vector of tract characteristics $X$ observed by the bidders, 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 in the sample (all auction formats) was 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. Figure 2.1 compares the distribution of bidders according to size (number of employees) in both types of auctions. Both distributions are remarkably similar, supporting the assumption that bidders in both auctions come from the same underlying population.

**Figure 2.1.** Distribution of bidder size. A comparison of first-price and ascending auctions.

2.2.1 Choosing $X$, the vector of conditioning variables

As I mentioned above, the test focuses on auctions with two bidders, so everything is done conditional on $n = 2$. More precisely, I condition on the specific configuration of types $((\ell, \ell), (\ell, s), (s, s))$ in the auction. 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 transaction price in both types of auctions is included in Table 2.3. The results there show that, once conditioning on number of bidders, the only auction characteristic that has explanatory power for transaction price in both types of auctions is reserve price (per volume of timber). For this reason and to mitigate the effect of the curse of dimensionality, reserve price is the only conditioning variable I include in $X$. Figure 2.2 shows the estimated density of reserve price in both auction formats. My test requires that the supports of the corresponding distributions of reserve price have a nonempty intersection, and the graph strongly indicates that this is the case. In fact, while the distributions are not necessarily the same (they are not assumed to be), the supports appear to be the same. The validity of that assumption appears to be supported by the data in my empirical illustration.

<table>
<thead>
<tr>
<th></th>
<th>Ascending auctions</th>
<th></th>
<th>Sealed auctions</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>S.E.</td>
<td>Coefficient</td>
<td>S.E.</td>
</tr>
<tr>
<td>Reserve price ($/mbf)</td>
<td>1.296***</td>
<td>(0.081)</td>
<td>1.121***</td>
<td>(0.036)</td>
</tr>
<tr>
<td>Volume (mbf)</td>
<td>-0.002</td>
<td>(0.001)</td>
<td>0.152</td>
<td>(0.209)</td>
</tr>
<tr>
<td>Herfindahl index</td>
<td>-0.023</td>
<td>(13.688)</td>
<td>6.824</td>
<td>(18.307)</td>
</tr>
<tr>
<td>Number of bidders</td>
<td>48.689***</td>
<td>(3.641)</td>
<td>39.400***</td>
<td>(3.534)</td>
</tr>
<tr>
<td>Road costs ($/mbf)</td>
<td>-8.532</td>
<td>(6.067)</td>
<td>15.151</td>
<td>(17.348)</td>
</tr>
<tr>
<td>Logging costs ($/mbf)</td>
<td>0.089</td>
<td>(0.544)</td>
<td>0.541**</td>
<td>(0.234)</td>
</tr>
<tr>
<td>Manufacturing costs ($/mbf)</td>
<td>-0.629</td>
<td>(0.687)</td>
<td>0.307</td>
<td>(0.407)</td>
</tr>
<tr>
<td>State</td>
<td>3.373</td>
<td>(6.556)</td>
<td>11.041</td>
<td>(9.621)</td>
</tr>
<tr>
<td>Density (mbf/acre)</td>
<td>0.007</td>
<td>(0.890)</td>
<td>-0.704</td>
<td>(0.948)</td>
</tr>
<tr>
<td>Contract duration ($/mbf)</td>
<td>5.443**</td>
<td>(2.578)</td>
<td>0.456</td>
<td>(1.355)</td>
</tr>
<tr>
<td>Controls (dummy variables)</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1923</td>
<td></td>
<td>902</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.3. Determinants of transaction price.

Note: The dependent variable is transaction price. The regression was estimated over the subsample of auctions that were run and had a transaction price reported. The estimation method was Nonlinear Least Squares with a sample-selection correction term that assumes a mean-zero normally distributed error term (whose variance was estimated along with the parameters reported). The sample sizes indicated therefore correspond to the auctions that were run and where a transaction price was recorded.
Figure 2.2. Density of the reserve price in ascending and sealed-bid auctions with n=2 players.

Kolmogorov-Smirnov test for the equality of distributions had a p-value = 0.0389.

<table>
<thead>
<tr>
<th></th>
<th>Ascending</th>
<th>Sealed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean($)</td>
<td>161.11</td>
<td>188.53</td>
</tr>
<tr>
<td>Std</td>
<td>155.69</td>
<td>161.72</td>
</tr>
</tbody>
</table>

2.3 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. Section 1.4.6 outlines how, 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).

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 (2017a), in my case the set of bidders in an auction is not fixed, i.e., different players may participate in a two bidders 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 higher 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 auctions. In Appendix B, I carry out the type of reduced-form exploratory tests for collusion proposed by Porter and Zona (1993). The analysis

---

1Evidence suggestive of collusion in sealed-bid timber auctions has been found for regions and time periods very different from mine: in the case of Canadian timber auctions for the period 1996-2000, in Schurter (2017b). In the case of USFS Region 1 (Idaho, Mondana and eastern Washington) for the time period 1983-1992, in Shen (2017).
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. To further mitigate the possible presence of collusion, I consider only auctions where bids were at least 1\% above the USFS reserve price.

2.4 Implementation of the test

I apply the test described in Section 1.3.2 for auctions with \( n = 2 \), conditional on \( X = \)reserve price (per unit of timber) and conditional on the configuration of bidder types observed \( ((\ell, \ell), (\ell, s), (s, s)) \). This allows me to test separately four potentially distinct values of \( k \):

- \( k_{ss} \): the number of iterated dominance steps small firms perform to bid against small firms.
- \( k_{sl} \): the number of iterated dominance steps small firms perform to bid against large firms.
- \( k_{ll} \): the number of iterated dominance steps large firms perform to bid against large firms.
- \( k_{ls} \): the number of iterated dominance steps large firms perform to bid against small firms.

I treat the reserve price as a continuous variable. So, the number of continuous 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 \([-1, 1] \).

\[
k(u) = (C_1(1 - u^2)^2 + C_2(1 - u^2)^4) 1(|u| \leq 1),
\]

where \( C_1 \) and \( C_2 \) are chosen in such way that \( \int_{-1}^{1} k(u)du = 1 \) and \( \int_{-1}^{1} u^j k(u)du = 0 \) 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 the nonparametric estimators are normalized to $0.5 \cdot \hat{\sigma}(X)$ (for $X$), $0.5 \cdot \hat{\sigma}(b)$ (for bids $b$) and $0.5 \cdot \hat{\sigma}(P)$ (for transaction price $P$), where $\hat{\sigma}(\cdot)$ is a standard deviation in the corresponding sample. The tuning parameters $\gamma_{L_1}$ and $\gamma_{L_2}$ are set to be $0.004^2$. The variance $\sigma_L^2$ described in Equation (1.11) is estimated by bootstrap with 1000 bootstrap draws. The tuning parameter $\lambda_L$ is set below machine precision. I consider values of $k$ ranging between 2 and 6 (values of $k$ larger than 6 produced estimated bounds that are computationally indistinguishable from those of $k = 6$). Table 2.4 summarizes the results of the test.

<table>
<thead>
<tr>
<th>$k$</th>
<th>$k_{ss}$</th>
<th>p-value</th>
<th>$k_{sf}$</th>
<th>p-value</th>
<th>$k_{\ell \ell}$</th>
<th>p-value</th>
<th>$k_{\ell s}$</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.341</td>
<td>0.090</td>
<td>1.118</td>
<td>0.132</td>
<td>0.421</td>
<td>0.337</td>
<td>1.845</td>
<td>0.033</td>
</tr>
<tr>
<td>3</td>
<td>1.439</td>
<td>0.075</td>
<td>2.095</td>
<td>0.018</td>
<td>0.399</td>
<td>0.345</td>
<td>1.851</td>
<td>0.032</td>
</tr>
<tr>
<td>4</td>
<td>1.460</td>
<td>0.072</td>
<td>2.096</td>
<td>0.018</td>
<td>0.394</td>
<td>0.347</td>
<td>1.867</td>
<td>0.031</td>
</tr>
<tr>
<td>5</td>
<td>1.463</td>
<td>0.072</td>
<td>2.096</td>
<td>0.018</td>
<td>0.394</td>
<td>0.347</td>
<td>1.869</td>
<td>0.031</td>
</tr>
<tr>
<td>6</td>
<td>1.463</td>
<td>0.072</td>
<td>2.096</td>
<td>0.018</td>
<td>0.394</td>
<td>0.347</td>
<td>1.869</td>
<td>0.031</td>
</tr>
</tbody>
</table>

At a 5% significance level, $k_{sf} \geq 3$ is rejected and $k_{\ell s} \geq 2$ is also rejected, while I cannot reject any $k \leq 6$ 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.

### 2.5 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

---

2The 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 tuning parameters $\gamma_{L_1}$ and $\gamma_{L_2}$
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. Symmetry yields a straightforward closed-form expression for the BNE bidding function (see Section 2.3 in Krishna (2010)). These are given by

\[ b_{BNE}^{\ell\ell}(v|x) = \frac{1}{F_{0,\ell}(v|x)} \int_v^\infty \frac{t \cdot f_{2\ell,P,X}(t|x)}{2 \cdot (1 - F_{0,\ell}(t|x))} dt, \]

\[ b_{BNE}^{ss}(v|x) = \frac{1}{F_{0,s}(v|x)} \int_v^\infty \frac{t \cdot f_{2s,P,X}(t|x)}{2 \cdot (1 - F_{0,s}(t|x))} dt, \]

where, as before, $\ell\ell$ and $ss$ refer to two-bidder auctions involving two large and two small bidders, respectively. These BNE bidding functions are strictly increasing and invertible in $v$ conditional on $x$. Let $v_{BNE}^{\ell\ell}(\cdot|x)$ and $v_{BNE}^{ss}(\cdot|x)$ denote their respective inverse functions. Let $G_{1}^{\ell\ell}(\cdot|x)$ and $G_{1}^{ss}(\cdot|x)$ denote the conditional cdf of bids in first-price auctions given $x$ in two-bidder auctions with two large and two small bidders, respectively. For a given $t \in (\bar{b}, \tilde{b})$ denote

\[ \phi_{BNE}^{\ell\ell}(t|x) = F_{0,\ell}(v_{BNE}^{\ell\ell}(t|x)|x) - G_{1}^{\ell\ell}(t|x), \]

\[ \phi_{BNE}^{ss}(t|x) = F_{0,s}(v_{BNE}^{ss}(t|x)|x) - G_{1}^{ss}(t|x). \]

One-sided violations of risk-neutral BNE can be tested via stochastic dominance restrictions, using the above functionals. I focus the test on auctions with two bidders of the same type because 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.

2.5.1 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 large and small bidders. Framing it as a stochastic-dominance restriction (like the test for $k$–rationalizability), this is a test of the null hypotheses $H_0 : \phi_{BNE}^{\ell\ell}(t|x) \leq 0$ for all $x \in \mathcal{X}$, $t \in \mathcal{B}(x)$ (for large bidders) and $H_0 : \phi_{BNE}^{ss}(t|x) \leq 0$ for all $x \in \mathcal{X}$, $t \in \mathcal{B}(x)$ (for small bidders). The conditions leading to Theorem 1.3.1 produce an
analogous result in this case. From here, 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) bids\(^3\).

### 2.5.2 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 $H_0: \phi_{BNE}^{\ell}(t|X) \geq 0$ for all $x \in X$, $t \in B(x)$ (for large bidders) and $H_0: \phi_{BNE}^{ss}(t|X) \geq 0$ for all $x \in X$, $t \in B(x)$ (for small bidders). 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 (see Section 4.1 in Krishna (2010)).

### 2.5.3 Two-sided BNE test

Testing the null hypothesis that all bidders play according to the risk-neutral BNE is equal to testing $H_0: \phi_{BNE}^{\ell}(t|X) = 0$ for all $x \in X$, $t \in B(x)$ (for large bidders) and $H_0: \phi_{BNE}^{ss}(t|X) = 0$ for all $x \in X$, $t \in B(x)$ (for small bidders). In this case, I use similar statistics. The only difference is in the step of constructing $\Lambda_{BNE}^{\ell}(x)$ and $\Lambda_{BNE}^{ss}(x)$ in (1.7):

\[
\Lambda_{BNE}^{\ell}(x) = \int_{t \in B(x)} \left[ (\phi_{BNE}^{\ell})_+ + (-\phi_{BNE}^{\ell})_+ \right] dQ(t|x),
\]

\[
\Lambda_{BNE}^{ss}(x) = \int_{t \in B(x)} \left[ (\phi_{BNE}^{ss})_+ + (-\phi_{BNE}^{ss})_+ \right] dQ(t|x).
\]

### 2.5.4 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 2.5. At a significance level of 5%, the results in Table 2.5 reject risk-neutral BNE in all kinds of two-bidder symmetric auctions. However, they suggest that violations

\(^3\)Since in my data set of ascending auctions, I have only auctions with competing behavior, I consistently estimate the distribution of private values and thus, the distribution of bids in BNE.
Table 2.5. Results of BNE tests in auctions with bidders of the same type.

<table>
<thead>
<tr>
<th></th>
<th>Small firms</th>
<th>Large firms</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_0 ): No overbidding above risk-neutral BNE</td>
<td>test-statistic</td>
<td>p-value</td>
</tr>
<tr>
<td></td>
<td>1.56</td>
<td>0.059</td>
</tr>
<tr>
<td>( H_0 ): Risk-averse BNE</td>
<td>test-statistic</td>
<td>p-value</td>
</tr>
<tr>
<td></td>
<td>2.23</td>
<td>0.013</td>
</tr>
<tr>
<td>( H_0 ): Risk-neutral BNE</td>
<td>test-statistic</td>
<td>p-value</td>
</tr>
<tr>
<td></td>
<td>2.41</td>
<td>0.008</td>
</tr>
</tbody>
</table>

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 \( k \geq 6 \).

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

The results from the \( k \)-rationalizability tests in Table 2.4 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 in those auctions. 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 \( b_{\ell,s}^t(\cdot|x) \) denote the bidding functions of large firms in two-bidder asymmetric auctions. Let

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

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

\[
F_0^{\ell,s}(\cdot|x) = \text{cdf of values for large bidders (assumed to be the same in first-price and in ascending auctions).}
\]

\[
F_{0,s}(\cdot|x) = \text{cdf of values for small bidders (assumed to be the same in first-price and in ascending auctions).}
\]

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

\[
\tilde{G}_{1,\ell}^{\ell,s}(t|x) = \hat{F}_{0,\ell}\left(\tilde{b}_{\ell,s}^{\ell,s-1}(t|x)\right) \tag{2.1}
\]

### 2.6.1 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 (2.1). It is a measure of the status-quo given by

\[
\pi_{\text{benchmark}}^{\ell}(v|x) = (v - b_{\ell,s}^{\ell,s}(v|x)) \cdot P(b_{\ell,s}^{\ell,s} \leq b_{\ell,s}^{\ell,s}(v|x)|x)
\]

\[
= (v - b_{\ell,s}^{\ell,s}(v|x)) \cdot G_{1,s}^{\ell,s}(b_{\ell,s}^{\ell,s}(v|x)|x).
\]

Benchmark expected payoffs for large bidders are therefore estimated as

\[
\tilde{\pi}_{\text{benchmark}}^{\ell}(v|x) = (v - \tilde{b}_{\ell,s}^{\ell,s}(v|x)) \cdot \tilde{G}_{1,s}^{\ell,s}(\tilde{b}_{\ell,s}^{\ell,s}(v|x)|x).
\]

**Boundary case.** The results in Table 2.4 indicate that, with positive probability, large bidders believe that small ones bid above the \( k = 2 \) rationalizable bounds. However, the results there indicate that small bidders in fact bid below this bound
w.p.1 (since \( k = 2 \) is not rejected for small bidders). Therefore large bidders overestimate, with positive probability, the bids from small bidders. Given the results, 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 \( B_{s_2}^{t,s}(\cdot|x) \) denote these bounds and let \( v_{s_2}^{t,s}(\cdot|x) \) 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

\[
\pi_{\text{boundary}}^{\ell}(v|x) = \max_{b \geq v} \left\{ (v - b) \cdot P(B_{s_2}^{t,s}(v_s|x) \leq b|x) \right\} \\
= \max_{b \geq v} \left\{ (v - b) \cdot F_{0,s}(v_{s_2}^{t,s}(b|x)|x) \right\}
\]

They can be estimated as

\[
\hat{\pi}_{\text{boundary}}^{\ell}(v|x) = \max_{b \geq v} \left\{ (v - b) \cdot \hat{F}_{0,s}(v_{s_2}^{t,s}(b|x)|x) \right\}.
\]

**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 \( b_{s_2}^{t,s}(\cdot|x) \). This measure of expected payoffs is given by

\[
\pi_{\text{optimal}}^{\ell}(v|x) = \max_{b \geq v} \left\{ (v - b) \cdot G_{1,s}(b|x) \right\}
\]

These payoffs can be estimated as

\[
\hat{\pi}_{\text{optimal}}^{\ell}(v|x) = \max_{b \geq v} \left\{ (v - b) \cdot \hat{G}_{1,s}(b|x) \right\}
\]

The goal in this counterfactual exercise is to estimate the economic impact for large bidders of having incorrect beliefs about small bidders. Notice that the three cases of expected payoff I consider correspond to three scenarios about beliefs:

**Benchmark:** Reflects large bidders’ current (incorrect) beliefs.

**Boundary:** The counterfactual case where large bidders adjust their beliefs to the most pessimistic beliefs consistent with the behavior of small bidders.

**Optimal:** The counterfactual case where large bidders have correct beliefs.
about the current bidding strategies of small bidders.

2.6.2 Results from comparing current expected payoffs against counterfactuals

Figure 2.3 compares expected payoffs, conditional on reserve price (i.e, integrating out $v$) for each one of the three cases (95% pointwise confidence bands are included in Figure 2.4). As one can see, the 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.

Figure 2.3. Expected payoffs conditional on reserve price.

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 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). Figure 2.3 does not appear to reveal any particular pattern.

I compare the difference in expected payoffs, conditional on reserve price, at the level of each individual auction. Tables 2.6-2.8 describe the median and the 75th 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.

Table 2.6. Counterfactual results: Difference in large bidders’ expected payoffs (in 2010 real dollars).

<table>
<thead>
<tr>
<th>All auctions with two asymmetric bidders</th>
<th>Difference in expected profits per-unit of timber</th>
<th>Difference in expected profits by total volume of timber</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Boundary minus benchmark</td>
<td>Optimal minus benchmark</td>
</tr>
<tr>
<td>Median</td>
<td>$692</td>
<td>$1,728</td>
</tr>
<tr>
<td>75th percentile</td>
<td>$6,919</td>
<td>$8,377</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Auctions with two asymmetric bidders won by large bidders</th>
<th>Difference in expected profits per-unit of timber</th>
<th>Difference in expected profits by total volume of timber</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Boundary minus benchmark</td>
<td>Optimal minus benchmark</td>
</tr>
<tr>
<td>Median</td>
<td>$324</td>
<td>$1,726</td>
</tr>
<tr>
<td>75th percentile</td>
<td>$7,089</td>
<td>$7,681</td>
</tr>
</tbody>
</table>

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

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 (i.e, moving from the benchmark to the boundary case) 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 and about 21 times the 75th percentile of advertised value.

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 75th percentile corresponds to an improvement of about 160% (approximately 7.5 times the 75th percentile of advertised value).

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 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 (auctions with the highest reserve-prices). Table 2.8 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 (according to total advertised value). Comparing the observed bids against the counterfactual optimal bids in the boundary case (where they best-respond to the highest possible bid that could have been submitted by their opponent according to the results), the median proportional amount of overbidding is 13% and the 75th percentile is 42%. Moving to the 10% most valuable tracts, these proportions are 42% and 45%, respectively.

### 2.7 Concluding remarks

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 rationalizable 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.
Figure 2.4. Expected payoffs conditional on reserve price, with 95% bootstrap point-wise confidence bands
Table 2.7. Counterfactual results: Percentage difference in large bidders' expected payoffs.

<table>
<thead>
<tr>
<th></th>
<th>Boundary Benchmark (%)</th>
<th>Optimal Benchmark (%)</th>
<th>Optimal Boundary (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All auctions with two asymmetric bidders</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>104%</td>
<td>112%</td>
<td>108%</td>
</tr>
<tr>
<td>75th percentile</td>
<td>258%</td>
<td>265%</td>
<td>124%</td>
</tr>
<tr>
<td><strong>Auctions with two asymmetric bidders won by large bidders</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>102%</td>
<td>105%</td>
<td>107%</td>
</tr>
<tr>
<td>75th percentile</td>
<td>150%</td>
<td>157%</td>
<td>122%</td>
</tr>
</tbody>
</table>

Table 2.8. Counterfactual results: Evidence of overbidding. Summary statistics for $\frac{b_{\text{benchmark}}(\cdot)}{b_{\text{boundary}}(\cdot)}$ (in % terms) in auctions won by large bidders.

<table>
<thead>
<tr>
<th></th>
<th>All auctions</th>
<th>High-value auctions</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Auctions with reserve price in upper 25th percentile</td>
<td>Auctions with reserve price in upper 10th percentile</td>
</tr>
<tr>
<td><strong>Median</strong></td>
<td>102%</td>
<td>113%</td>
<td>141%</td>
</tr>
<tr>
<td><strong>75th percentile</strong></td>
<td>127%</td>
<td>142%</td>
<td>145%</td>
</tr>
</tbody>
</table>
Chapter 3  
Testing functional inequalities conditional on an estimated index

3.1 Introduction

The approach proposed in this chapter can be applied in cases where testable implications of econometric models are in the form of conditional functional inequalities. An example of such economic problem was considered in the previous chapters of this dissertation. A test for the rationalizable behavior of bidders in first-price auctions has the form of conditional inequality. In Guerre et al. (2000), the authors develop a testable implication of the Bayesian-Nash equilibrium in first-price auctions which can be rewritten in the form of conditional inequality. The nonparametric approach to this question described in previous chapters or in Aradillas-Lopez et al. (2016) allows to condition on any number of observables both continuous or discrete. But in many applied problems, the researcher can have a rich collection of observables to choose from. The “curse of dimensionality” grows very fast with the number of conditioning observables. If the number of observations is relatively small, it may lead to poor finite sample properties. For example, in Chapter 2, I describe a timber auction data set that contains more than 10 observable characteristics of the auction. There, I selected only one observable characteristic as conditioning variable. One possible solution to reduce the dimensionality of the conditioning variable can be estimated index/indexes.
3.2 Model

Assume a testable implication of the model is

$$\mathbb{E}[Y|X] \leq 0, \text{ a.s. in } X$$  \hfill (3.1)

Let \((X_i, Y_i)_{i=1}^n\) be iid. Denote \(T \equiv X'\beta^*\) and \(f_T(\cdot)\) is pdf of the random variable \(T\), where \(\beta^*\) is an unknown vector of parameters. Assume there exists an estimator\(^1\)

$$\hat{\beta} = \beta^* + \frac{1}{n} \sum_{j=1}^n \psi_{\beta_j} + o_p(n^{-\delta}),$$  \hfill (3.2)

where \(\delta \geq \frac{1}{2}\) and \(\psi_{\beta_j}\) is a known transformation of the data (influence functions) with the following properties:

(i) \(\mathbb{E}[\psi_{\beta_j}^\beta] = 0\) for all \(j = 1, \ldots, n\)

(ii) Let \(\mathbb{E}[(\psi_{\beta_j}^\beta)^2] \equiv \sigma_{j}^{2,\beta}\) and \(\lim_{j \to \infty} \sigma_{j}^{2,\beta} = \sigma^{2,\beta} > 0\).

Denote,

$$\mu(t) \equiv \mathbb{E} [Y|X'\beta^* = t] = \mathbb{E} [Y|T = t].$$ \hfill (2)

The dimension of conditioning variable \(X\) depends on the real application and can be very large. The performance of nonparametric methods for analyzing the inequality (3.1) can be poor in small samples if the number of conditioning variable is large. This “curse of dimensionality” becomes severe very fast. So, instead of testing the initial inequality (3.1), I consider the following inequality:

$$\mathbb{E}[Y|X'\beta^*] \leq 0, \text{ a.s. in } X'\beta^*$$  \hfill (3.3)

**Remark 3.2.1** If the exclusion restriction \(Y|X \sim Y|X'\beta^*\) holds, then testing the initial inequality (3.1) is equivalent to testing the inequality (3.3).

\(^1\hat{\beta}\) can be estimated from an auxiliary data. The following results stay the same if the sizes of data sets grow proportionally.

\(^2\)The whole analysis can be extended to testing the inequality of a smooth transformation of a linear index that can be nonparametrically estimated.
The exclusion restriction in Remark 3.2.1 is testable. For example, in Fan and Li (1996) the authors describe a nonparametric test that can be used to test this exclusion restriction.

Note that the “curse of dimensionality” can be solved not only by an aggregate single linear index, but also by any number of smooth indexes. So, $X'\beta^*$ can be replaced with multiple linear indexes: $X'_1\beta^*_1, \ldots, X'_p\beta^*_p$ or with smooth known functions $g_1(X), \ldots, g_p(X)$. The following results can also be extended to the case of nonparametrically estimated index function $\hat{g}(X)$ under suitable bandwidth convergence rate and smoothness conditions, similar to Aradillas-Lopez et al. (2007). However, if $g(\cdot)$ is a function from $\mathbb{R}^c$ to $\mathbb{R}$ where $c \equiv \dim (X)$, then the “curse of dimensionality” problem shifts from estimation $E[Y|X]$ nonparametrically to estimation of $g(X)$. One example when nonparametric estimation of the index function can be useful is when it takes the form of $g(X'\beta^*)$. This problem is studied, for example, in Powell et al. (1989).

**Remark 3.2.2** The inequality (3.1) implies the inequality (3.3) for any joint distribution of $Y$ and $X^3$.

The statement of the previous remark (3.2.2) follows from the law of iterated expectation. $E[Y|X'] = E[E[Y|X, X'\beta^*]|X'] = E[E[Y|X]|X'\beta^*]$. So, if $E[Y|X] \leq 0$ a.s. in $X$ then $E[Y|X'] \leq 0$ a.s. in $X'\beta^*$. Thus, the inequality (3.3) is an implication of the inequality (3.1).

Therefore, even if exclusion restriction described in Remark 3.2.1 is not satisfied, testing (3.3) is still a valid procedure but it can be conservative.

### 3.3 Testing procedure

I focus on testing the inequality (3.3). Note that testing inequality (3.3) is equivalent to testing $E[Y|X'\beta^*]f(X'\beta^*) \leq 0$ a.s. in $X'\beta^*$ for any strictly positive function $f(\cdot)$. It is convenient to choose the probability distribution function $f_T(\cdot)$ as weighting function because it simplifies estimator of $R(t)$ since I do not need to estimate the denominator of the conditional expectation. Denote

$$R(t) \equiv \mu(t)f_T(t) = E[Y|X'\beta^* = t]f_T(t) = E[Y|T = t]f_T(t).$$

\(^3\)The equivalent statement is a violation of the inequality (3.3) necessarily implies a violation of (3.1), but the converse is not necessarily true.
The main hypothesis I consider in the chapter is

\[ H_0 : R(t) \leq 0, \text{ a.s. } t \in \mathcal{T}, \]  

(3.4)

where \( \mathcal{T} \subseteq \text{supp}(T) \) is a fixed testing range.

### 3.3.1 An estimator for \( R(t) \)

I use a Nadaraya-Watson type estimator for \( R(t) \).

\[
\hat{R}(t) \equiv \frac{1}{nh_n} \sum_{i=1}^{n} Y_i K \left( \frac{X_i' \hat{\beta} - t}{h_n} \right),
\]

(3.5)

where \( K(\cdot) \) is a kernel function and \( h_n \) is a bandwidth.

First, I study the properties of the proposed estimator of \( R(t) \) and then the properties of the statistic for testing the inequality (3.4), which uses \( \hat{R}(t) \). The following assumption on the kernel function is sufficient to derive the limit distribution of the estimator \( \hat{R}(t) \). Note that since the kernel function is a researcher’s choice, the assumption on the kernel function impose no assumption on the model and data.

**Assumption 3.3.1** The kernel \( K(\cdot) \) is a bounded function, symmetric around zero and it has support over a compact interval \([-k_S, k_S]\). It is a bias-reducing kernel of order \( M \) (restrictions on \( M \) will be described below).

\[
\int_{-k_S}^{k_S} K(u)du = 1, \quad \int_{-k_S}^{k_S} u^j K(u)du = 0, \quad j = 1, \ldots, M - 1, \quad \int_{-k_S}^{k_S} |u|^M K(u)du < \infty,
\]

where \( K(\cdot) \) is twice differentiable almost everywhere with bounded second derivative \( K''(\cdot) \).

Estimation with kernels requires smoothness conditions:

**Assumption 3.3.2** The following functions \( \mu(t) = \mathbb{E}[Y|X'\beta^* = t] \), \( \sigma_{Y|T}(t) \equiv \mathbb{E}[Y^2|X'\beta^* = t] \), \( \nu(t) \equiv \mathbb{E}[YX'|X'\beta^* = t] \) and \( \eta(t) \equiv \mathbb{E}[YY'X'|X'\beta^* = t] \) are \( M \)-times continuously differentiable with bounded derivatives and \( \mathbb{E}[|YX'\psi^\beta|] < \infty \).

The last set of assumptions needed to define limiting properties of \( \hat{R}(t) \) related to the rate of convergence of bandwidth \( h_n \) and properties of \( \hat{\beta} \) and order of kernel \( M \).
Assumption 3.3.3 The bandwidth $h_n$, the order of kernel $M$ and $\delta$ satisfy relative rate of convergence as $n \to \infty$:

(i) $n^{\frac{1}{2}}h_n^M \to 0$

(ii) $n^\delta h_n^2 \to \infty$

(iii) $nh_n^3 \to \infty$

For example, if $h_n = n^{-\alpha}$ for some $\alpha > 0$, the conditions for the relative rate of convergence are the following:

\[
\begin{align*}
-\frac{1}{2} - \alpha M &< -\frac{1}{2}, \\
-1 + \frac{3}{2} \alpha &< -\frac{1}{2}, \\
-\frac{3}{2} + 2\alpha &< -\frac{1}{2}, \\
-\alpha M &< -\frac{1}{2}, \\
-\delta - \frac{1}{2} + 2\alpha &< -\frac{1}{2}, \\
-\frac{3}{2} + 3\alpha &< -\frac{1}{2}.
\end{align*}
\]

$\iff$

\[
\begin{align*}
\alpha &< \frac{1}{3}, \\
\alpha &< \frac{1}{2}\delta, \\
\alpha &> \frac{1}{2M}.
\end{align*}
\]

From Assumption (3.3.3) and definition of $\hat{\beta}$ (3.2), it follows that

\[
\begin{align*}
\delta &\geq \frac{1}{2}, \\
\delta &> \frac{1}{M}, \\
M &\geq 2.
\end{align*}
\]

So, there is a trade-off between the rate of convergence of the remaining term in the estimator $\beta^*$ (3.2) and the smoothness requirements together with the order of the kernel. $\delta$ can be equal to $\frac{1}{2}$ if $M \geq 3$.

The majority of regular extremum estimators satisfy the assumption (3.3.3) with $\delta = \frac{1}{2}$ and require $M$ to be strictly larger than 2. A list of examples in parametric models can be found in Newey and McFadden (1994) including MLE, NLS, GMM and CMD. In the context of semiparametric index models, an example of an estimator that satisfies assumption (3.3.3) is a semiparametric estimator in Powell et al. (1989).
Theorem 3.3.1 If assumption (3.3.1), (3.3.2) and (3.3.3) are satisfied then

\[ \hat{R}(t) - R(t) = \frac{1}{n} \sum_{i=1}^{n} \psi_i^{R,1}(\psi_i^\beta, t) + \frac{1}{nh_n} \sum_{i=1}^{n} \psi_i^{R,2}(Y_i, X_i, t; t_n) + \xi_n^R(t), \]  

(3.6)

where by construction \( \mathbb{E} \left[ \psi_i^{R,1}(\psi_i^\beta, t) \right] = \mathbb{E} \left[ \psi_i^{R,2}(Y_i, X_i, t; t_n) \right] = 0 \) and \( \sup_{t \in \mathcal{T}} |\xi_n^R(t)| = O_p \left( n^{-\frac{1}{2}} \right) \), for some \( \epsilon > 0 \).

Note that the first two terms in (3.6) have different rates of convergence: \( O_p(n^{-\frac{1}{2}}) \) and \( o_p(n^{-\frac{1}{2}} h_n^{-1}) \) respectively. If I only want to analyze properties the estimator \( \hat{R}(t) \), the first term will be a remaining term. In the next chapter, I show how to test the inequality (3.4) and the properties of the estimator \( \hat{R}(t) \) are key for those of the statistics used to test it. After averaging, both terms produce the leading term of the final statistic, which converges at a parametric rate. Thus, even though the two leading terms on the right hand side of (3.6) converge at different rates, they are both relevant to the asymptotic properties of the proposed test-statistic.

3.4 The test-statistics

My goal is to test the hypothesis that \( R(t) \leq 0 \) over the testing range \( \mathcal{T} \), i.e.

\[ H_0 : R(t) \leq 0, \text{ a.s in } \mathcal{T} \quad \text{vs.} \quad H_1 : R(t) > 0, \text{ w.p } 0 \]  

(3.7)

In this section I construct a statistic to test (3.7) and describe its limiting properties.

Let \( Q \) be some measure on the testing range \( \mathcal{T} \). Then testing hypothesis \( H_0 \) in (3.7) is equivalent to testing

\[ H_0 : S \equiv \int \max\{R(t), 0\} dQ(t) = 0, \quad \text{vs.} \quad H_1 : \int \max\{R(t), 0\} dQ(t) > 0 \]

To estimate the statistic \( S \), I plug in the estimator \( \hat{R}(t) \) (3.5) described in the previous section and take the maximum between this estimator and a tuning parameter.
that converges to zero to facilitate the limiting properties of the estimator. This is the same testing approach used in previous chapters.

Thus, I estimate $S$ with

$$
\hat{S} = \int \hat{R}(t) 1\{\hat{R}(t) \geq -b_n\} dQ(t).
$$

To analyze the limiting properties of $\hat{S}$, I need to impose assumption on the influence functions $\psi_{R,1}$ and $\psi_{R,2}$, rate of convergence of a new tuning parameter $b_n$ and behavior of $R(t)$ around zero.

**Assumption 3.4.1** The following two classes of functions

$$
R_1 = \{\Psi^{R,1}(\cdot; t) : t \in \mathcal{T}\} \quad \text{and} \quad R_2 = \{\Psi^{R,2}(\cdot, \cdot; t, h_n) : t \in \mathcal{T}, h_n > 0\}
$$

are Euclidean with a constant envelope

**Assumption 3.4.2** There exists some $D > 0$ such that

$$
\lim_{b \to 0} \frac{P_Q\{-2b \leq R(t) < 0\}}{b} \leq D
$$

**Assumption 3.4.3** The rate of convergence of the tuning parameter $b_n$ is such that $b_n n^{1/2} \to \infty$ and $b_n n^{1/2} h_n \to \infty$.

**Theorem 3.4.1** If assumptions (3.4.1), (3.4.2) and (3.4.3) are satisfied then the estimator $\hat{S}$ has the following linear representation:

$$
\hat{S} = S + \frac{1}{n} \sum_{i=1}^{n} \psi^S_i(Y_i, X_i, \psi^\beta_i; h_n) + \xi_n^S,
$$

where

$$
\psi^S_i(Y_i, X_i, \psi^\beta_i; h_n) = \int \left[ \psi_{R,1}^i(\psi^\beta_i, t) + \frac{1}{h_n} \psi_{R,2}^i(Y_i, X_i, t; h_n) \right] 1\{R(t) \geq 0\} dQ(t)
$$

and $|\xi_n^S| = O_p(n^{-1/2-\epsilon})$ for some $\epsilon > 0$. The influence function satisfies the following conditions:

(i) $E[\psi^S_i(Y_i, X_i, \psi^\beta_i; h_n)] = 0$
(ii) If \( R(t) < 0 \) for almost all \( t \in \mathcal{T} \) then \( \psi^S_i(Y_i, X_i, \psi^\beta; h_n) = 0 \) w.p.1

(iii) Let \( \mathbb{V}[\psi^S_i(Y_i, X_i, \psi^\beta; h_n)] = \sigma^2_{S,n} \). If \( R(t) \geq 0 \) with positive probability in \( \mathcal{T} \) then

\[
\lim_{n \to \infty} \sigma^2_{S,n} = \sigma^2_S > 0.
\]

**Proof:** The first and the second properties of the influence function follow directly from its definition.

To show the form of the influence function consider the limiting properties of the estimator:

\[
\hat{S} - S = \int \hat{R}(t)1\{\hat{R}(t) \geq -b_n\} dQ(t) - \int R(t)1\{R(t) \geq 0\} dQ(t) = \quad (3.8)
\]

\[
= \int \left( \hat{R}(t) - R(t) \right) 1\{R(t) \geq 0\} dQ(t) + \int R(t) \left( 1\{\hat{R}(t) \geq -b_n\} - 1\{R(t) \geq 0\} \right) dQ(t) +
\]

\[
+ \int \left( \hat{R}(t) - R(t) \right) \left( 1\{\hat{R}(t) \geq -b_n\} - 1\{R(t) \geq 0\} \right) dQ(t)
\]

The first term in \( (3.8) \) is the leading term. To study its limiting properties plug in the result from theorem (3.3.1): (3.8):

\[
\int \left( \hat{R}(t) - R(t) \right) 1\{R(t) \leq 0\} dQ(t) =
\]

\[
= \int \left( \frac{1}{n} \sum_{i=1}^n \psi^R_{i}^1(\psi^\beta_i, t) + \frac{1}{nh_n} \psi^R_{i}^2(Y_i, X_i, t; h_n) + \xi^R_n(t) \right) 1\{R(t) \geq 0\} dQ(t) =
\]

\[
= \frac{1}{n} \sum_{i=1}^n \int \left[ \psi^R_{i}^1(\psi^\beta_i, t) + \frac{1}{h_n} \psi^R_{i}^2(Y_i, X_i, t; h_n) \right] 1\{R(t) \geq 0\} dQ(t) + \int \xi^R_n(t) 1\{R(t) \geq 0\} dQ(t)
\]

Denote

\[
\psi^S_i(Y_i, X_i, \psi^\beta; h_n) = \int \left[ \psi^R_{i}^1(\psi^\beta_i, t) + \frac{1}{h_n} \psi^R_{i}^2(Y_i, X_i, t; h_n) \right] 1\{R(t) \geq 0\} dQ(t),
\]

\[
\xi^S_n = \int \xi^R_n(t) 1\{R(t) \geq 0\} dQ(t).
\]

Thus,

\[
\int \left( \hat{R}(t) - R(t) \right) 1\{R(t) \leq 0\} dQ(t) = \frac{1}{n} \sum_{i=1}^n \psi^S_i(Y_i, X_i, \psi^\beta; h_n) + \xi^S_n.
\]

82
Now, I show that \( \mathbb{E}[\psi_i^S(Y_i, X_i, \psi_i^\beta; h_n)] = 0 \) and \( \nabla[\psi_i^S(Y_i, X_i, \psi_i^\beta; h_n)] < \infty \) and \( |\xi_n^{S,1}| = O_p\left(n^{-\frac{1}{2} - \epsilon}\right) \), for some \( \epsilon > 0 \).

\[
\mathbb{E}[\psi_i^S(Y_i, X_i, \psi_i^\beta; h_n)] = \mathbb{E} \left[ \int \left[ \psi_i^{R,1}(\psi_i^\beta, t) + \frac{1}{h_n} \psi_i^{R,2}(Y_i, X_i, t; h_n) \right] 1\{R(t) \geq 0\} dQ(t) \right] = \\
\int \left( \mathbb{E} \left[ \psi_i^{R,1}(\psi_i^\beta, t) \right] + \frac{1}{h_n} \mathbb{E} \left[ \psi_i^{R,2}(Y_i, X_i, t; h_n) \right] \right) 1\{R(t) \geq 0\} dQ(t) = 0.
\]

\[
\nabla[\psi_i^S(Y_i, X_i, \psi_i^\beta; h_n)] = \mathbb{E} \left[ \left( \psi_i^S(Y_i, X_i, \psi_i^\beta; h_n) \right)^2 \right] = \\
= \mathbb{E} \left[ \left( \int \psi_i^{R,1}(\psi_i^\beta, t) 1\{R(t) \geq 0\} dQ(t) + \frac{1}{h_n} \int \psi_i^{R,2}(Y_i, X_i, t; h_n) 1\{R(t) \geq 0\} dQ(t) \right)^2 \right]
\]

The first integral does not depend on \( h_n \) and

\[
\int \psi_i^{R,1}(\psi_i^\beta, t) 1\{R(t) \geq 0\} dQ(t) = - \int \frac{n-1}{n} (\nu(t) f_T(t))' \psi_i^\beta 1\{R(t) \geq 0\} dQ(t) = \\
= - \frac{n-1}{n} \psi_i^\beta \int (\nu(t) f_T(t))' 1\{R(t) \geq 0\} dQ(t).
\]

Denote \( A \equiv \int (\nu(t) f_T(t))' 1\{R(t) \geq 0\} dQ(t) \) and assume it is bounded.

The second integral:

\[
\frac{1}{h_n} \int \psi_i^{R,2}(Y_i, X_i, t; h_n) 1\{R(t) \geq 0\} dQ(t) = \\
= \frac{1}{h_n} \int \left( -Y_i K \left( \frac{X'_i \beta^* - t}{h_n} \right) + \mathbb{E} \left[ Y K \left( \frac{X'_i \beta^* - t}{h_n} \right) \right] \right) 1\{R(t) \geq 0\} dQ(t)
\]

Choose \( Q(t) \) such that there exists \( q(t) = Q'(t) \) and change variables \( t = h_n v + X'_i \beta^* \)

\[
\frac{1}{h_n} \int Y_i K \left( \frac{X'_i \beta^* - t}{h_n} \right) 1\{R(t) \geq 0\} dQ(t) = \\
= \int Y_i K(v) 1 \{ R(h_n v + X'_i \beta^*) \geq 0 \} q(h_n v + X'_i \beta^*) dt.
\]

If \( \mathbb{E}[|Y|] \) and \( \nabla[Y] \) exist, the kernel \( K(\cdot) \) and the measure \( Q(\cdot) \) can be chosen such that

\[
\nabla \left[ \frac{1}{h_n} \int Y K \left( \frac{X'_i \beta^* - t}{h_n} \right) 1\{R(t) \geq 0\} dQ(t) \right] = O_p(1), \quad (3.10)
\]

83
$$E \left[ \psi_i \beta Y_i \int K(v) \{ R(h_nv + X_i'\beta^*) \geq 0 \} q(h_nv + X_i'\beta^*)dt \right] = O_p(1). \quad (3.11)$$

So, $V \left[ \psi_i^S(Y_i, X_i, \psi_i^\beta, t; h_n) \right] = O_p(1)$.

Since $\sup_{t \in T} |\xi_n^R(t)| = O_p \left(n^{-\frac{1}{2} - \epsilon}\right)$, for some $\epsilon > 0$ and $|\int 1\{ R(t) \geq 0 \}dQ(t)| \geq 1$, $|\xi_n^{S,1}| = O_p \left(n^{-\frac{1}{2} - \epsilon}\right)$, for some $\epsilon > 0$. Therefore, the leading term in (3.8) has a parametric rate of convergence $O_p(n^{-\frac{1}{2}})$. The second and the third terms in (3.8) vanish in probability faster than $O_p(n^{-\frac{1}{2}})$. All details of the proof can be found in the Appendix (C.2).

The third property of the influence function follows from the (3.9), (3.10) and (3.11).

From here on, the testing procedure of the inequality (3.4) would follow all the same steps from the section (1.3.2.5).

### 3.5 Conclusion

In this chapter I consider a method of addressing the “curse of dimensionality” problem in testing an inequality of an estimated functional conditional on many observables. The whole procedure consists of three steps. The first step is focused on reducing the dimensionality of the conditioning variable by estimating a single linear index or an index function. It can be estimated using the main data set or an auxiliary data. In the second step, an unknown functional conditional on the index estimated in the first step is estimated nonparametrically, pointwise, using kernels. Even though the nonparametric part of this estimator converges slower than the parametric estimator of the linear index, both terms will be relevant asymptotically in the distribution of the final statistic. The statistic is of the form of a weighted average of the pointwise estimated functional. The resulting statistic converges at a parametric rate and both parametric and nonparametric leading terms from the second step produce the leading term of the final statistic. The method described here can be extended to the case of many linear indices or nonlinear index functions and to the inequality of unknown but identifiable functionals that satisfy smoothness conditions as well as the technical conditions required in this chapter.
Appendix A
Econometric Appendix

A.1 Proof of Theorem 1.3.1. Preliminary results

Before proving Theorem 1.3.1, I will first characterize the relevant asymptotic properties of the nonparametric estimators \( \hat{G}_1(t|x) \) and \( \hat{F}_0(\tau_k(t|x)|x) \). The testing range sets \( \mathcal{X} \) and \( \mathcal{B} \) are those described in Assumption 1.3.4.

A.1.1 An asymptotic linear representation result for \( \hat{G}_1(t|x) - G_1(t|x) \)

Take the marginal kernel \( k \) described in Assumption 1.3.9. 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 \( \sup_{u \in \mathbb{R}} |\delta(u)| \leq \int_{-S}^{S} |k(\psi)| d\psi < \infty \). Now, for bid \( b^j \) in first-price auction \( j \) and a fixed \( b \) let

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

Note that

\[
m(b^j, b; h_{b,L_1}) = \frac{1}{h_{b,L_1}} \int_{-\infty}^{b} k \left( \frac{b - b^j}{h_{b,L_1}} \right) db,
\]

and under the conditions in Assumptions 3.3.2 and 1.3.9,

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

As in the main text, group all the bids submitted in the \( j^{th} \) first-price auction as \( b_j \),
and in an auction with \( n \) bids submitted, define

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

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

\[
\hat{R}_1(b|x) = \frac{1}{L_1 h_{x,L_1}^e} \sum_{j=1}^{L_1} \mathcal{H}(X_j - x; h_{x,L_1}) \cdot \overline{m}_n \left( b_j, b; h_{b,L_1} \right).
\]

The estimator \( \hat{G}_1(b|x) \) described in Section 1.3.2.3 can be expressed as

\[
\hat{G}_1(b|x) = \frac{\hat{R}_1(b|x)}{\hat{f}_{1,x}(x)}.
\]

Fix \( x \) and, once again, let \( n \) denote the element in \( x \) corresponding to number of participating bidders. Fix \( b \), and define

\[
\phi_{R_1}^{G_1} \left( b_j, X_j, b, x; h_{b,L_1}, h_{x,L_1} \right) = \mathcal{H}(X_j - x; h_{x,L_1}) \cdot \overline{m}_n \left( b_j, b; h_{b,L_1} \right) - \mathbb{E} \left[ \mathcal{H}(X_j - x; h_{x,L_1}) \cdot \overline{m}_n \left( b_j, b; h_{b,L_1} \right) \right],
\]

\[
\phi_{f_1,x}^{G_1} \left( X_j, x; h_{x,L_1} \right) = \mathcal{H}(X_j - x; h_{x,L_1}) - \mathbb{E} \left[ \mathcal{H}(X_j - x; h_{x,L_1}) \right],
\]

\[
\phi_{G_1}^{G_1} \left( b_j, X_j, b, x; h_{b,L_1}, h_{x,L_1} \right) = \frac{1}{\hat{f}_{1,x}(x)} \cdot \phi_{R_1}^{G_1} \left( b_j, X_j, b, x; h_{b,L_1}, h_{x,L_1} \right) - \frac{G_1(b|x)}{\hat{f}_{1,x}(x)} \cdot \phi_{f_1,x}^{G_1} \left( X_j, x; h_{x,L_1} \right).
\]

Note that \( \mathbb{E} \left[ \phi_{G_1}^{G_1} \left( b_j, X_j, b, x; h_{b,L_1}, h_{x,L_1} \right) \right] = 0 \). Let \( \mathcal{B} \) be any compact subset of \( \text{int}(\overline{[b, \underline{b}]}) \). A second-order approximation along with the conditions in Assumptions 3.3.2 and 1.3.9 produce the following result,

\[
\hat{G}_1(b|x) - G_1(b|x) = \frac{1}{L_1 h_{x,L_1}^e} \sum_{j=1}^{L_1} \phi_{G_1}^{G_1} \left( b_j, X_j, b, x; h_{b,L_1}, h_{x,L_1} \right) + \xi_{L_1}^{G_1}(b, x),
\]

where

\[
\sup_{b \in \mathcal{B}, x \in \mathcal{X}} \left| \phi_{L_1}^{G_1}(b, x) \right| = O \left( h_{b,L_1} \right) + O \left( h_{x,L_1}^M \right) = O \left( L_1^{-1/2-\epsilon} \right) \quad \text{for some} \ \epsilon > 0.
\]

(A.1)

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

Take the marginal kernel $k$ described in Assumption 1.3.9. 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

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

Note that

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

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

$$\hat{F}_{2,P\mid X}(v|x) = \int_{-\infty}^{v} \hat{f}_{2,P\mid X}(t|x) dt = \frac{1}{f_{2,X}(x)} \cdot \frac{1}{L_2 h_{x,L}} \sum_{i=1}^{L_2} m(P, v; h_{p,L}) \cdot H(X_i - x; h_{x,L}).$$

For a given $(v, x)$, define

$$\varphi_{2,P\mid X}^R(P; X_i, v, x; h_{p,L}, h_{x,L}) =$$

$$m(P; v; h_{p,L}) \cdot H(X_i - x; h_{x,L}) - E \left[ m(P; v; h_{p,L}) \cdot H(X_i - x; h_{x,L}) \right],$$

$$\varphi_{2,x}^f (X_i, x; h_{x,L}) =$$

$$H(X_i - x; h_{x,L}) - E \left[ H(X_i - x; h_{x,L}) \right],$$

$$\varphi_{2,P\mid X}^F(P; X_i, v; h; h_{p,L}, h_{x,L}) =$$

$$\frac{1}{f_{2,x}(x)} \cdot \varphi_{2,P\mid X}^R(P; X_i, v; x; h_{p,L}, h_{x,L}) - \frac{F_{2,P\mid X}(v|x)}{f_{2,x}(x)} \cdot \varphi_{2,x}^f (X_i, x; h_{x,L}).$$

Note that $E \left[ \varphi_{2,P\mid X}^F(P; X_i, v; h; h_{p,L}, h_{x,L}) \right] = 0$. Let $\mathcal{V}$ be any compact subset of $int([\underline{v}, \overline{v}])$. A second-order approximation along with the conditions in Assumptions
3.3.2 and 1.3.9 yield

\[ \hat{F}_{2,P|X}(v|x) - F_{2,P|X}(v|x) = \frac{1}{L_2 h_{x,L_2}^c} \sum_{i=1}^{L_2} \varphi^{F_{2,P|X}}(P_i, X_i, v, x; h_{p,L_2}, h_{x,L_2}) + \xi^{F_{2,P|X}}_{L_2}(v, x), \]

where 

\[ \sup_{v \in \mathcal{V}} \left| \xi^{F_{2,P|X}}_{L_2}(v, x) \right| = O(h_{p,L_2}) + O(h_{x,L_2}^{M}) = O(L_2^{-1/2-\epsilon}) \quad \text{for some } \epsilon > 0. \]

\[ \sup_{v \in \mathcal{V}} \left| \hat{F}_{2,P|X}(v|x) - F_{2,P|X}(v|x) \right| = O_p \left( \frac{1}{\sqrt{L_2^{-1/2-\epsilon}}} \right) \quad \forall \delta > 0. \]  \hspace{1cm} (A.2)

Therefore, from our bandwidth convergence restrictions we get

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

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

\[ \Gamma^{(1)}(F_{2,P|X}(v|x)) = \frac{1}{n(n-1)F_{2,P|X}(v|x)n^{-2}(1-F_{2,P|X}(v|x))}, \]

\[ \varphi^{F_0}(P_i, X_i, v, x; h_{p,L_2}, h_{x,L_2}) = \Gamma^{(1)}(F_{2,P|X}(v|x)) \cdot \varphi^{F_{2,P|X}}(P_i, X_i, v, x; h_{p,L_2}, h_{x,L_2}). \]  \hspace{1cm} (A.3)

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

\[ \hat{F}_0(v|x) - F_0(v|x) = \frac{1}{L_2 h_{x,L_2}^c} \sum_{i=1}^{L_2} \varphi^{F_0}(P_i, X_i, v, x; h_{p,L_2}, h_{x,L_2}) + \xi^{F_0}_{L_2}(v, x), \]

where 

\[ \sup_{v \in \mathcal{V}} \left| \xi^{F_0}_{L_2}(v, x) \right| = O_p \left( L_2^{-1/2-\epsilon} \right) \quad \text{for some } \epsilon > 0. \]  \hspace{1cm} (A.4)

Asymptotic properties of \( \hat{F}_0 \left( \hat{v}_k(t|x) \right) - F_0(\bar{v}_k(t|x)|x) \) 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 $\nabla_b \pi = \frac{\partial \pi}{\partial b}$. We have,

$$\nabla_b \pi = -F_n^{-1}(b|x) + (n-1) \cdot \frac{(v-b)F_n^{-2}(b|x)}{\Gamma^{(1)}(F_2,P|X(b|x)) | f_2,P|X(b|x)|}.$$

Let

$$b^*_k(v|x) = \arg\max_{b \geq 0} \pi_k(v,b,x).$$

Under our assumptions, $b^*_k(v|x)$ is characterized by the first-order conditions

$$\nabla_b \pi_k(v,b^*_k(v|x),x) = 0. \quad (A.5)$$

By definition of the $k-$rationalizable bounds,

$$\overline{B}_k(v|x) = v - \pi_k(v,b^*_k(v|x),x).$$

My estimator for $\overline{B}_k(v|x)$ is

$$\hat{\overline{B}}_k(v|x) = v - \hat{\pi}_k(v,b^*_k(v|x),x),$$

where

$$\hat{b}^*_k(v|x) = \arg\max_{b \geq 0} \hat{\pi}_k(v,b,x),$$

with

$$\hat{\pi}_k(v,b,x) = (v-b)\hat{F}_n^{-1}(b|x).$$

By the design of my estimator $\hat{F}_n(b|x)$, the derivative $\nabla_b \hat{\pi}_k = \frac{\partial \hat{\pi}_k}{\partial b}$ has the exact sample-analog structure of $\nabla_b \pi_k = \frac{\partial \pi_k}{\partial b}$. Namely,

$$\nabla_b \hat{\pi}_k(v,b,x) = -\hat{F}_n^{-1}(b|x) + (n-1) \cdot \frac{(v-b)\hat{F}_n^{-2}(b|x)}{\Gamma^{(1)}(\hat{F}_2,P|X(b|x)) | \hat{f}_2,P|X(b|x)|}.$$

and $\hat{b}^*_k(v|x)$ satisfies the sample-analog first order conditions

$$\nabla_b \hat{\pi}_k(v,\hat{b}^*_k(v|x),x) = 0. \quad (A.6)$$

The primary goal is to characterize the asymptotic properties of $\hat{\overline{B}}_k(v|x) - \overline{B}_k(v|x)$. Let $\mathcal{V}$ be any compact subset of $\text{int}([\underline{v}, \bar{v}])$ and let $\mathcal{B}$ be any compact subset of $\text{int}([\underline{b}, \bar{b}])$. Using our previous results and bandwidth convergence conditions in As-
sumption 1.3.9, we have

\[
\sup_{v \in \mathcal{V}, b \in \mathcal{B}, x \in \mathcal{X}} |\nabla b \hat{\pi}_{k=2}(v, b, x) - \nabla b \pi_{k=2}(v, b, x)| = O_p \left( L_2^{-1/4-\epsilon/2} \right)
\]

for some \( \epsilon > 0 \). Combining this with (A.5) and (A.6), we also obtain \( \sup_{v \in \mathcal{V}, x \in \mathcal{X}} \left| \hat{b}_{k=2}^*(v|x) - b_{k=2}^*(v|x) \right| = O_p \left( L_2^{-1/4-\epsilon/2} \right) \) for some \( \epsilon > 0 \).

From a second-order approximation we get

\[
\hat{\pi}_{k=2}(v, b_{k=2}^*(v|x), x) = \hat{\pi}_{k=2}(v, \tilde{b}_{k=2}^*(v|x), x) + \nabla \hat{\pi}_{k=2}(v, \tilde{b}_{k=2}^*(v|x), x) \cdot \left( b_{k=2}^*(v|x) - \hat{b}_{k=2}^*(v|x) \right) = 0 \quad \text{from (A.6)}
\]

\[
+ O_p \left( \left| \hat{b}_{k=2}^*(v|x) - b_{k=2}^*(v|x) \right|^2 \right).
\]

Therefore,

\[
\sup_{v \in \mathcal{V}, x \in \mathcal{X}} \left| \hat{\pi}_{k=2}(v, \tilde{b}_{k=2}^*(v|x), x) - \hat{\pi}_{k=2}(v, b_{k=2}^*(v|x), x) \right| = O_p \left( L_2^{-1/2-\epsilon} \right) \quad \text{for some} \quad \epsilon > 0.
\]

(A.7)

And from here,

\[
\hat{\mathcal{B}}_{k=2}(v|x) = \mathcal{B}_{k=2}(v|x) + \left( \hat{\pi}_{k=2}(v, b_{k=2}^*(v|x), x) - \pi_{k=2}(v, b_{k=2}^*(v|x), x) \right) + \hat{\varphi}_{L_2}^{k=2}(v, x),
\]

where \( \sup_{v \in \mathcal{V}} \left| \hat{\varphi}_{L_2}^{k=2}(v, x) \right| = O_p \left( L_2^{-1/2-\epsilon} \right) \) for some \( \epsilon > 0 \).

(A.8)

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

\[
\varphi^{F_{k=2}}(P_1, X_i, v, x; h_{p,L_2}, h_{x,L_2}) = (n - 1) \cdot F_0^{n-2} \left( b_{k=2}^*(v|x) | x \right) \cdot \left( v - b_{k=2}^*(v|x) \right) \cdot \varphi^{F_0}(P_1, X_i, b_{k=2}^*(v|x), x; h_{p,L_2}, h_{x,L_2})
\]

(A.9)
Note that $E \left[ \varphi^k (P_i, X_i, v, x; h_{p,L_2}, h_{x,L_2}) \right] = 0$. Combining (A.4) and (A.8) obtains,

$$
\hat{B}_{k=2}(v|x) - B_{k=2}(v|x) = \frac{1}{L_2 h_{x,L_2}} \sum_{i=1}^{L_2} \varphi^k (P_i, X_i, v, x; h_{p,L_2}, h_{x,L_2}) + \xi_{L_2}^k(v,x),
$$

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

\begin{equation}
(A.10)
\end{equation}

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

\[ v - (v - b_{k=2}^*(v|x)) F_0^{n-1} (b_{k=2}^*(v|x)|x) = t, \]

\[ v - (v - \hat{b}_{k=2}^*(v|x)) \hat{F}_0^{n-1} (\hat{b}_{k=2}^*(v|x)|x) = t. \]

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

\[ \nabla_v \bar{B}_{k=2}(v|x) = 1 - F_0^{n-1} (b_{k=2}^*(v|x)|x), \]

\[ \nabla_v \hat{B}_{k=2}(v|x) = 1 - \hat{F}_0^{n-1} (\hat{b}_{k=2}^*(v|x)|x). \]

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

\[ \varphi_{\bar{v}_{k=2}} (P_i, X_i, t, x; h_{p,L_2}, h_{x,L_2}) = \]

\[ - \frac{1}{1 - F_0^{n-1} (b_{k=2}^*(\bar{v}_{k=2}(t|x)|x)|x)} \cdot \varphi_{\bar{B}_{k=2}} (P_i, X_i, \bar{v}_{k=2}(t|x), x; h_{p,L_2}, h_{x,L_2}) \]

\begin{equation}
(A.11)
\end{equation}
Note that $E \left[ \varphi^{F_{k=2}}(P_t, X_i, t, x; h_{p,L_2}, h_{x,L_2}) \right] = 0$. Our previous results yield,

$$
\widehat{\varphi}_{k=2}(t|x) - \overline{\varphi}_{k=2}(t|x) = \frac{1}{L_2 h_{x,L_2}} \sum_{i=1}^{L_2} \varphi^{F_{k=2}}(P_t, X_i, t, x; h_{p,L_2}, h_{x,L_2}) + \xi^{F_{k=2}}_{L_2}(t, x),
$$

where $\sup_{t \leq t' \leq t} \xi^{F_{k=2}}_{L_2}(t, x) = O_p \left( L_2^{-1/2-\epsilon} \right)$ for some $\epsilon > 0$. \hfill (A.12)

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

$$
\varphi^{F_{0,k=2}}(P_t, X_i, t, x; h_{p,L_2}, h_{x,L_2}) = \\
\varphi^{F_0}(P_t, X_i, \pi_{k=2}(t|x), x; h_{p,L_2}, h_{x,L_2}) \\
+ \Gamma(1) \left( F_{2,p,X}(\pi_{k=2}(t|x)|x) \right) \cdot f_{2,p,X}(\pi_{k=2}(t|x)|x) \cdot \varphi^{F_{k=2}}(P_t, X_i, t, x; h_{p,L_2}, h_{x,L_2}). \hfill (A.13)
$$

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

$$
\widehat{F}_0 \left( \pi_{k=2}(t|x)|x \right) - F_0 \left( \pi_{k=2}(t|x)|x \right) = \frac{1}{L_2 h_{x,L_2}} \sum_{i=1}^{L_2} \varphi^{F_{0,k=2}}(P_t, X_i, t, x; h_{p,L_2}, h_{x,L_2}) + \xi^{F_{0,k=2}}_{L_2}(t, x),
$$

where $\sup_{t \leq t' \leq t} \xi^{F_{0,k=2}}_{L_2}(t, x) = O_p \left( L_2^{-1/2-\epsilon} \right)$ for some $\epsilon > 0$. \hfill (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 $\widehat{F}_0(\pi_k(t|x)|x) - F_0(\pi_k(t|x)|x)$ for $k \geq 3$**

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

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

$$
\widehat{\pi}_k(v, b, x) = (v - b) \widehat{F}_0^{n-1}(\widehat{\pi}_{k-1}(b|x)|x),
$$

with

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

92
Suppose the estimator $\hat{v}_{k-1}$ is such that

$$
\sup_{b \in \mathcal{B}} \sup_{x \in \mathcal{X}} \left| \hat{v}_{k-1}(b|x) - \bar{v}_{k-1}(b|x) \right|^2 = O_p \left( L_2^{-1/2-\epsilon} \right) \text{ for some } \epsilon > 0,
$$

note that this is true for $k = 3$ by the result in (A.12). As we defined in Assumption 1.3.4, let

$$
\mathcal{V}_k = \{ v : v = \pi_k(b|x) \text{ for some } b \in \mathcal{B} \text{ and } x \in \mathcal{X} \}.
$$

Then, by the conditions described in Assumption 1.3.4, a second-order approximation coupled with the first-order conditions of $b^*_k$ and $\bar{b}^*_k$ and the Envelope Theorem yield a generalization of the result in (A.7). Namely,

$$
\sup_{v \in \mathcal{V}_k} \sup_{x \in \mathcal{X}} \left| \hat{\pi}_k \left( v, \hat{b}^*_k(v|x), x \right) - \hat{\pi}_k \left( v, b^*_k(v|x), x \right) \right| = O_p \left( L_2^{-1/2-\epsilon} \right) \text{ for some } \epsilon > 0. \quad (A.15)
$$

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

$$
\hat{B}_k(v|x) = B_k(v|x) + \left( \hat{\pi}_k \left( v, b^*_k(v|x), x \right) - \pi_k \left( v, b^*_k(v|x), x \right) \right) + \varphi_{L_2}^k(v, x),
$$

where

$$
\sup_{v \in \mathcal{V}_k} \sup_{x \in \mathcal{X}} \left| \varphi_{L_2}^k(v, x) \right| = O_p \left( L_2^{-1/2-\epsilon} \right) \text{ for some } \epsilon > 0. \quad (A.16)
$$

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

$$
\hat{v}_{k-1}(t|x) - \bar{v}_{k-1}(t|x) = \frac{1}{L_2h_{x,L_2}^*} \sum_{i=1}^{L_2} \varphi^{k-1} \left( P_i, X_i, t, x; h_{p,L_2}, h_{x,L_2} \right) + \xi_{k-1}^{L_2}(t, x),
$$

where

$$
\sup_{t \in \mathcal{B}} \sup_{x \in \mathcal{X}} \left| \xi_{k-1}^{L_2}(t, x) \right| = O_p \left( L_2^{-1/2-\epsilon} \right) \text{ for some } \epsilon > 0. \quad (A.17)
$$

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

93
\[ \pi_k(v, b^*_k(v|x), x) \]. This is given by

\[
\begin{align*}
\hat{\pi}_k(v, b^*_k(v|x), x) - \pi_k(v, b^*_k(v|x), x) & = (v - b^*_k(v|x)) \times \left[ \frac{F_{\nu_0}^{n-1}(\nu_{k-1}(b^*_k(v|x)|x))}{\nu_{k-1}(b^*_k(v|x)|x)} - F_{\nu_0}^{n-1}(\nu_{k-1}(b^*_k(v|x)|x)) \right] \\
& = (v - b^*_k(v|x)) \times \left\{ \frac{F_{\nu_0}^{n-1}(\nu_{k-1}(b^*_k(v|x)|x))}{\nu_{k-1}(b^*_k(v|x)|x)} - F_{\nu_0}^{n-1}(\nu_{k-1}(b^*_k(v|x)|x)) \right\} \\
& \quad + \left[ \frac{F_{\nu_0}^{n-1}(\nu_{k-1}(b^*_k(v|x)|x))}{\nu_{k-1}(b^*_k(v|x)|x)} - F_{\nu_0}^{n-1}(\nu_{k-1}(b^*_k(v|x)|x)) \right] \\
\end{align*}
\]

(A.18)

Notice that the term \( \left[ \frac{F_{\nu_0}^{n-1}(\nu_{k-1}(b^*_k(v|x)|x))}{\nu_{k-1}(b^*_k(v|x)|x)} - F_{\nu_0}^{n-1}(\nu_{k-1}(b^*_k(v|x)|x)) \right] \) was equal to zero in the case \( k = 2 \) because \( \nu_1(t|x) = \nu_{k-1}(t|x) = t \), since \( \nu_{k-1}(v|x) = \nu_{k-1}(v|x)|x \) (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 \( \varphi^{F_0} \) and \( \varphi^{\nu_{k-1}} \) 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

\[
\varphi^{\nu_k} (P_i, X_i, v, x; h_{p,L_2}, h_{x,L_2}) = \\
(n - 1) \cdot F_{\nu_0}^{n-2}(\nu_{k-1}(b^*_k(v|x)|x)) \cdot (v - b^*_k(v|x)) \times \left\{ \varphi^{F_0} (P_i, X_i, \nu_{k-1}(b^*_k(v|x)|x), x; h_{p,L_2}, h_{x,L_2}) \\
+ \Gamma^{(1)}(F_{2,P|X}(\nu_{k-1}(b^*_k(v|x)|x))) \cdot f_{2,P|X}(\nu_{k-1}(b^*_k(v|x)|x)) \cdot \varphi^{\nu_{k-1}} (P_i, X_i, b^*_k(v|x), x; h_{p,L_2}, h_{x,L_2}) \right\}
\]

(A.19)

Note that \( E \left[ \varphi^{\nu_k} (P_i, X_i, v, x; h_{p,L_2}, h_{x,L_2}) \right] = 0 \), 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, \( \nu_1(t|x) = \nu_1(t|x) = t \) (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),

\[
\begin{align*}
\hat{\nu}_k(v|x) & = \nu_k(v|x) + \frac{1}{L_2 h_{x,L_2}^c} \sum_{i=1}^{L_2} \varphi^{\nu_k} (P_i, X_i, v, x; h_{p,L_2}, h_{x,L_2}) + \xi^{\nu_k}_{L_2}(v, x), \\
\end{align*}
\]

where \( \sup_{v \in \nu_k} \xi^{\nu_k}_{L_2}(v, x) = O_p \left( L_2^{-1/2-\epsilon} \right) \) for some \( \epsilon > 0 \).
By the Envelope Theorem and the first-order conditions satisfied by \( b_k^*(v|x) \) and \( \hat{b}_k(x) \), from the definition of \( \overline{B}_k(v|x) \) and \( \hat{B}_k(v|x) \) we obtain \( \nabla_v \hat{B}_k(v|x) = 1 - \hat{F}_{0}^{n-1}(b_k^*(v|x)|x) \) and \( \nabla_v \hat{B}_k(v|x) = 1 - \hat{F}_{0}^{n-1}(b_k^*(v|x)|x) \). Once again, let \( \tau_k(\cdot|x) \) and \( \hat{\tau}_k(\cdot|x) \) denote the inverse functions of \( \overline{B}_k(\cdot|x) \) and \( \hat{B}_k(\cdot|x) \), respectively. By construction, for any \( t \) we must have \( \hat{B}_k(\hat{\tau}_k(t|x)|x) = \overline{B}_k(\tau_k(t|x)|x) \) (since both expressions are equal to \( t \), by definition). From here, we get

\[
\hat{B}_k(\hat{\tau}_k(t|x)|x) - \hat{B}_k(\tau_k(t|x)|x) = \overline{B}_k(\hat{\tau}_k(t|x)|x) - \overline{B}_k(\tau_k(t|x)|x).
\]

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

\[
\varphi^\tau_k (P_i, X_i, t, x; h_{p,L_2}, h_{x,L_2}) = \frac{1}{1 - \hat{F}_{0}^{n-1}(b_k^*(v(t|x)|x))} \cdot \varphi^\overline{\tau}_k (P_i, X_i, \tau_k(t|x), x; h_{p,L_2}, h_{x,L_2}) \tag{A.21}
\]

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

\[
\hat{\tau}_k(t|x) - \tau_k(t|x) = \frac{1}{L_2 h_{x,L_2}} \sum_{i=1}^{L_2} \varphi^\tau_k (P_i, X_i, t, x; h_{p,L_2}, h_{x,L_2}) + \xi^{\tau}_L(t, x), \tag{A.22}
\]

where \( \sup_{t \in \mathbb{R}, x \in \mathcal{X}} |\xi^{\tau}_L(t, x)| = \text{O}_p \left( L_2^{-1/2-\epsilon} \right) \) for some \( \epsilon > 0 \).

From here we can study the properties of \( \hat{F}_{0} \left( \hat{\tau}_k(t|x)|x \right) - \hat{F}_0(\hat{\tau}_k(t|x)|x) \) straightforwardly. Fix \( t \) and \( x \). Let \( \varphi^{\hat{F}_0} \) and \( \varphi^{\tau_k} \) be as described in (A.3) and (A.21) and define

\[
\varphi^{F_{0,k}}(P_i, X_i, t, x; h_{p,L_2}, h_{x,L_2}) = \varphi^{\hat{F}_0} (P_i, X_i, \overline{\tau}_k(t|x), x; h_{p,L_2}, h_{x,L_2}) \tag{A.23}
\]

\[
+ \Gamma^{(1)} \left( F_{2,P|X}(\overline{\tau}_k(t|x)|x) \right) \cdot f_{2,P|X}(\overline{\tau}_k(t|x)|x) \cdot \varphi^{\tau_k} (P_i, X_i, t, x; h_{p,L_2}, h_{x,L_2}).
\]

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

\[ \begin{align*} 
\hat{F}_0 \left( \overline{v}_k(t|x) \right) - F_0 \left( \overline{v}_k(t|x) \right) &= \frac{1}{L_2 h_{x,L_2}^c} \sum_{i=1}^{L_2} \varphi_{F_0,k}^L (P_i, X_i, t, x; h_{p,L_2}, h_{x,L_2}) + \xi_{L_2}^{F_0,k} (t, x), \\
\text{where } \sup_{t \in \mathcal{B}} \sup_{x \in \mathcal{X}} \left| \xi_{L_2}^{F_0,k} (t, x) \right| &= O_p \left( L_2^{-1/2} \right)^{1/2} \text{ for some } \epsilon > 0. 
\end{align*} \]

(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 \( \hat{\phi}_k(t|x) - \phi_k(t|x) \)

Combining our previous characterizations we obtain the result that will allow us to prove Theorem 1.3.1. Recall that

\[ \hat{\phi}_k(t|x) = \hat{F}_0 \left( \overline{v}_k(t|x) \right) - \hat{G}_1(t|x). \]

Also recall that I denoted \( L \equiv L_1 + L_2 \) as the combined sample sizes for first-price and ascending auctions. Let \( \varphi^{G_1} \) and \( \varphi^{F_0,k} \) be as described in (A.1) and (A.24), respectively. From the results obtained there, we get

\[ \hat{\phi}_k(t|x) - \phi_k(t|x) = \frac{1}{L_2 h_{x,L_2}^c} \sum_{i=1}^{L_2} \varphi_{F_0,k}^L (P_i, X_i, t, x; h_{p,L_2}, h_{x,L_2}) \]

\[ - \frac{1}{L_1 h_{x,L_1}^c} \sum_{j=1}^{L_1} \varphi^{G_1} (b_j, X_j, t, x; h_{b,L_1}, h_{x,L_1}) + \xi_{L_2}^{F_0,k} (t, x) - \xi_{L_1}^{G_1} (t, x), \]

where

\[ \sup_{t \in \mathcal{B}} \sup_{x \in \mathcal{X}} \left| \xi_{L_2}^{F_0,k} (t, x) \right| = O_p \left( L_2^{-1/2} \right)^{1/2} \text{ for some } \epsilon > 0, \]

\[ \sup_{t \in \mathcal{B}} \sup_{x \in \mathcal{X}} \left| \xi_{L_1}^{G_1} (t, x) \right| = O_p \left( L_1^{-1/2} \right)^{1/2}. \]

(A.25)
A.2 Proof of Theorem 1.3.1. Final steps

Recall that

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

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

\[ \frac{1}{L_1} \sum_{i=1}^{L_1} \hat{\Lambda}_{k,1}(X_i)W_{\mathcal{X}}(X_i) \]

\[ = \frac{1}{L_1} \sum_{i=1}^{L_1} \int_{t \in \mathcal{B}(X_i)} \hat{\phi}_k(t|X_i) \cdot 1 \left[ \hat{\phi}_k(t|X_i) \geq -\gamma_{L_1} \right] dQ(t|X_i) \cdot W_{\mathcal{X}}(X_i) \]

\[ = \frac{1}{L_1} \sum_{i=1}^{L_1} \int_{t \in \mathcal{B}(X_i)} \hat{\phi}_k(t|X_i) \cdot 1 \left[ \phi_k(t|X_i) \geq 0 \right] dQ(t|X_i)W_{\mathcal{X}}(X_i) \]

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

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

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

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

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

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

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

(A.26)
Recall that $\int_{t \in \mathcal{B}(x)} dQ(t|x) = 1$ for all $x$. Also note that $|\phi_k(t|x)| \leq 1$ for all $t, x$. Therefore,

$$
\frac{1}{L_1} \sum_{i=1}^{L_1} \int_{t \in \mathcal{B}(x_i)} \varepsilon_{k,1}(t|X_i) dQ(t|X_i) \leq \frac{1}{L_1} \sum_{i=1}^{L_1} \int_{t \in \mathcal{B}(x_i)} |\varepsilon_{k,1}(t|X_i)| dQ(t|X_i)
$$

$$
\leq 2 \times 1 \left[ \sup_{t \in \mathcal{B}, x \in \mathcal{X}} \left| \hat{\phi}_k(t|x) - \phi_k(t|x) \right| \geq \gamma_{L_1} \right]
$$

$$
+ 2 \gamma_{L_1} \times \sup_{t \in \mathcal{B}} \frac{1}{L_1} \sum_{i=1}^{L_1} 1 [-2 \gamma_{L_1} \leq \phi_k(t|X_i) < 0] \cdot W_{\mathcal{X}}(X_i)
$$

$$
+ 2 \cdot \sup_{t \in \mathcal{B}, x \in \mathcal{X}} \left| \hat{\phi}_k(t|x) - \phi_k(t|x) \right| \cdot 1 \left[ \sup_{t \in \mathcal{B}, x \in \mathcal{X}} \left| \hat{\phi}_k(t|x) - \phi_k(t|x) \right| \geq \gamma_{L_1} \right]
$$

$$
+ \sup_{t \in \mathcal{B}, x \in \mathcal{X}} \left| \hat{\phi}_k(t|x) - \phi_k(t|x) \right| \cdot \sup_{t \in \mathcal{B}} \frac{1}{L_1} \sum_{i=1}^{L_1} 1 [-2 \gamma_{L_1} \leq \phi_k(t|X_i) < 0] W_{\mathcal{X}}(X_i)
$$

(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 1.3.4 and 3.3.2, combined with the bounded-variation properties of the kernel described in Assumption 1.3.9 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,

$$
\mathcal{F}_k = \left\{ \phi^{F_{0,k}}(\cdot, \cdot, t, x; h_1, h_2): t \in \mathcal{B}, x \in \mathcal{X}, h_1 > 0, h_2 > 0 \right\},
$$

$$
\mathcal{G} = \left\{ \phi^{G_{1,k}}(\cdot, \cdot, t, x; h_1, h_2): t \in \mathcal{B}, x \in \mathcal{X}, h_1 > 0, h_2 > 0 \right\}
$$

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

$$
\nu_{L_1}(t, x) = \frac{1}{L_1} \sum_{i=1}^{L_1} \phi^{G_1}(b_i, X_i, t, x; h_{b,L_1}, h_{x,L_1}),
$$

$$
\nu_{L_2}^k(t, x) = \frac{1}{L_2} \sum_{i=1}^{L_2} \phi^{F_{0,k}}(P_i, X_i, t, x; h_{b,L_1}, h_{x,L_1}).
$$

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 $D$ such
that, for any $\delta > 0$,

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

Next, note that we can express

$$
\hat{\phi}_k(t|x) - \phi_k(t|x) = \frac{1}{h_{x,L_2}} \cdot \nu_{L_2}(t, x) - \frac{1}{h_{x,L_1}} \cdot \xi_{L_1}^{F_0,k}(t, x) - \xi_{L_1}^{G_1}(t, x),
$$

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

$$
\Pr \left[ \sup_{t \in \mathcal{B}} \left| \xi_{L_1}^{G_1}(t, x) \right| > \delta \right] \leq \frac{C}{\delta \cdot L_1^{1/2+\epsilon}}, \quad \text{and} \quad \Pr \left[ \sup_{t \in \mathcal{B}} \left| \xi_{L_2}^{F_0,k}(t, x) \right| > \delta \right] \leq \frac{C}{\delta \cdot L_2^{1/2+\epsilon}},
$$

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

$$
\Pr \left[ \sup_{t \in \mathcal{B}} \left| \hat{\phi}_k(t|x) - \phi_k(t|x) \right| > \gamma_{L_1} \right] \\ \leq \Pr \left[ \sup_{t \in \mathcal{B}} \left| \frac{1}{h_{x,L_1}} \nu_{L_1}(t, x) \right| > \frac{\gamma_{L_1}}{4} \right] + \Pr \left[ \sup_{t \in \mathcal{B}} \left| \frac{1}{h_{x,L_2}} \nu_{L_2}(t, x) \right| > \frac{\gamma_{L_1}}{4} \right] \\ + \Pr \left[ \sup_{t \in \mathcal{B}} \left| \xi_{L_1}^{G_1}(t, x) \right| > \frac{\gamma_{L_1}}{4} \right] + \Pr \left[ \sup_{t \in \mathcal{B}} \left| \xi_{L_2}^{F_0,k}(t, x) \right| > \frac{\gamma_{L_1}}{4} \right] \\ \leq \frac{4D}{\gamma_{L_1} h_{x,L_1} L_1^{1/2}} + \frac{4D}{\gamma_{L_1} h_{x,L_2} L_2^{1/2}} + \frac{4C}{\gamma_{L_1} L_1^{1/2+\epsilon}} + \frac{4C}{\gamma_{L_1} L_2^{1/2+\epsilon}} \rightarrow 0 \text{ as } L_1 \rightarrow \infty, \ L_2 \rightarrow \infty.
$$

Where the last result follows from the bandwidth convergence conditions in Assumption 1.3.9. Now let us go back to the indicator function $1 \left[ \sup_{t \in \mathcal{B}} \left| \hat{\phi}_k(t|x) - \phi_k(t|x) \right| \geq \gamma_{L_1} \right]$.
which appears in (A.27). Fix any $\Delta > 0$ and $\delta > 0$. Then,

$$
Pr \left[ L_{-1}^{\Delta} \left[ \sup_{t \in \mathcal{B}, x \in \mathcal{X}} |\hat{\phi}_k(t|x) - \phi_k(t|x)| \geq \gamma_{L_1} \right] > \delta \right] \leq Pr \left[ \sup_{t \in \mathcal{B}, x \in \mathcal{X}} |\hat{\phi}_k(t|x) - \phi_k(t|x)| > \gamma_{L_1} \right] \to 0.
$$

Therefore,

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

Lastly, the above conditions also imply

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

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

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

By Assumption 1.3.8, Lemma 2.12 in Pakes and Pollard (1989) and the Main Corollary in Sherman (1994), this process satisfies $\sup_{t \in \mathcal{B}} |\nu_{L_1}^{\phi_k} (t)| = O_p \left( L_1^{-1/2} \right)$. Combined with Assumption 1.3.7, this obtains

$$
\sup_{t \in \mathcal{B}} \frac{1}{L_1} \sum_{i=1}^{L_1} \left( 1 \left[ 2\gamma_{L_1} \leq \phi_k(t|X_i) < 0 \right] \cdot W_x (X_i) \right)
$$

$$
\leq \sup_{t \in \mathcal{B}} E \left[ 1 \left[ 2\gamma_{L_1} \leq \phi_k(t|X_i) < 0 \right] \cdot W_x (X_i) \right] + \sup_{t \in \mathcal{B}} \left| \nu_{L_1}^{\phi_k} (t) \right|
$$

$$
= O (\gamma_{L_1}) + O_p \left( L_1^{-1/2} \right) = O_p (\gamma_{L_1}). \tag{A.30}
$$
Plugging in the results from (A.28), (A.29) and (A.30) into (A.26) and (A.27), for any \( \Delta > 0 \) we obtain

\[
\frac{1}{L_1} \sum_{i=1}^{L_1} \hat{\Lambda}_{k,1}(X_i) W_{\mathcal{X}}(X_i) = \frac{1}{L_1} \sum_{i=1}^{L_1} \int_{t \in \mathcal{B}(X_i)} \tilde{\phi}_k(t|X_i) \cdot 1 [\phi_k(t|X_i) \geq 0] dQ(t|X_i) W_{\mathcal{X}}(X_i) + \hat{\delta}_{k,1},
\]

where \( \hat{\delta}_{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^c_{x,L_1}} \right) + O_p \left( \frac{\gamma_{L_1}}{L_2^{1/2} h^c_{x,L_2}} \right) \]

\[
= O_p \left( L_1^{-1/2-\epsilon} \right) + O_p \left( L_2^{-1/2-\epsilon} \right) \quad \text{for some } \epsilon > 0. \tag{A.31}
\]

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

\[
\frac{1}{L_2} \sum_{i=1}^{L_2} \hat{\Lambda}_{k,2}(X_i) W_{\mathcal{X}}(X_i) = \frac{1}{L_2} \sum_{i=1}^{L_2} \int_{t \in \mathcal{B}(X_i)} \tilde{\phi}_k(t|X_i) \cdot 1 [\phi_k(t|X_i) \geq 0] dQ(t|X_i) W_{\mathcal{X}}(X_i) + \hat{\delta}_{k,2},
\]

where \( \hat{\delta}_{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^c_{x,L_1}} \right) + O_p \left( \frac{\gamma_{L_2}}{L_2^{1/2} h^c_{x,L_2}} \right) \]

\[
= O_p \left( L_1^{-1/2-\epsilon} \right) + O_p \left( L_2^{-1/2-\epsilon} \right) \quad \text{for some } \epsilon > 0. \tag{A.32}
\]

Recall that \( \Lambda_k(x) = \int_{x \in \mathcal{B}(x)} (\phi_k(t|x)) dQ(t|x) \). Combining (A.25) and (A.31),

\[
\frac{1}{L_1} \sum_{i=1}^{L_1} \hat{\Lambda}_{k,1}(X_i) W_{\mathcal{X}}(X_i) = E_{1,X} \left[ \Lambda_k(X_i) W_{\mathcal{X}}(X_i) \right]
\]

\[
+ \frac{1}{L_1} \sum_{i=1}^{L_1} \left( \Lambda_k(X_i) W_{\mathcal{X}}(X_i) - E_{1,X} \left[ \Lambda_k(X_i) W_{\mathcal{X}}(X_i) \right] \right)
\]

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

\[
- \frac{1}{L_2^2 h^c_{x,L_2}} \sum_{i=1}^{L_1} \sum_{m=1}^{L_2} \int_{t \in \mathcal{B}(X_i)} \varphi^{G_{1}}(b_{m, X_m, t, X_i; h_{b,L_1}, h_{x,L_1}}) 1 [\phi_k(t|X_i) \geq 0] dQ(t|X_i) W_{\mathcal{X}}(X_i) + \tilde{\omega}_{k,1},
\]

where \( \tilde{\omega}_{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. \tag{A.33}
\]
Similarly, combining (A.25) and (A.32),

\[
\frac{1}{L_2} \sum_{i=1}^{L_2} \hat{A}_{k,2}(X_i) W_{X_i}(X_i) = E_{2,X} \left[ A_k(X_i) W_{X_i}(X_i) \right] \\
+ \frac{1}{L_2} \sum_{i=1}^{L_2} \left( A_k(X_i) W_{X_i}(X_i) - E_{2,X} \left[ A_k(X_i) W_{X_i}(X_i) \right] \right) \\
+ \frac{1}{L_2 h^{c}_{x,L_2}} \sum_{i=1}^{L_2} \sum_{\ell=1}^{L_2} \int_{t \in \mathcal{B}(X_i)} \varphi^{F_{0,k}}(P_t, X_t, t, X_i; h_{b,L_2}, h_{x,L_2}) 1 \left[ \phi_k(t|X_i) \geq 0 \right] dQ(t|X_i) \cdot W_{X_i}(X_i) \\
- \frac{1}{L_2 L_1 h^{c}_{x,L_1}} \sum_{i=1}^{L_2} \sum_{m=1}^{L_1} \int_{t \in \mathcal{B}(X_i)} \varphi^{G_1}(b_m, X_m, t, X_i; h_{b,L_1}, h_{x,L_1}) 1 \left[ \phi_k(t|X_i) \geq 0 \right] dQ(t|X_i) W_{X_i}(X_i) + \tilde{\omega}_{k,2},
\]

where \( \tilde{\omega}_{k,2} = O_p \left( L_1^{-1/2-\epsilon} \right) + O_p \left( L_2^{-1/2-\epsilon} \right) \) for some \( \epsilon > 0 \).

The result in Theorem 1.3.1 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

\[
U_{L_1}^a = \frac{1}{L_2 h^{c}_{x,L_1}} \sum_{i=1}^{L_1} \sum_{m=1}^{L_1} \int_{t \in \mathcal{B}(X_i)} \varphi^{G_1}(b_m, X_m, t, X_i; h_{b,L_1}, h_{x,L_1}) 1 \left[ \phi_k(t|X_i) \geq 0 \right] dQ(t|X_i) W_{X_i}(X_i)
\]

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

\[
\Xi_1^{G_1}(b, x, X^c; h_{b,L_1}, h_{x,L_1}) = \\
E_{1,X^c|X^c} \left[ \int_{t \in \mathcal{B}(X)} \varphi^{G_1}(b, x, t, X; h_{b,L_1}, h_{x,L_1}) 1 \left[ \phi_k(t|X) \geq 0 \right] dQ(t|X) W_{X}(X) \right] X^c,
\]

\[
\lambda_1^{G_1}(b_i, X_i; h_{b,L_1}, h_{x,L_1}) = \\
\int_{\psi \in [-S,S]^c} \Xi_1^{G_1}(b_i, X_i, h_{x,L_1} \psi + X_i^c; h_{b,L_1}, h_{x,L_1}) f_{1,X^c}(h_{x,L_1} \psi + X_i^c) d\psi.
\]
Notice that $E_{1,(b,X)} \left[ \lambda^G_1(b_i, X_i; h_{tb, L_1}, h_{x, L_1}) \right] = 0$ since, by iterated expectations,

$$E_{1,(b,X)} \left[ \lambda^G_1(b_i, X_i; h_{tb, L_1}, h_{x, L_1}) \right] =$$

$$E_{1,X} \left[ E_{1,(b,X)} \left[ \int_{t \in \mathcal{G}(X_i)} \varphi^{G_1}(b_m, X_m, t, X_i; h_{tb, L_1}, h_{x, L_1}) \mid X_i \right] 1 \left[ \phi(t | X_i) \geq 0 \right] dQ(t | X_i) W_X(X_i) \right] = 0.$$  

The Hoeffding decomposition of $U^p_{L_1}$ yields the following result,

$$U^p_{L_1} = \frac{1}{L_1} \sum_{i=1}^{L_1} \lambda^G_1(b_i, X_i; h_{tb, L_1}, h_{x, L_1}) + \tilde{\varphi}^a,$$

where $\tilde{\varphi}^a = O_p \left( \frac{1}{L_1 h^c_{x, L_1}} \right) = O_p \left( L_1^{-1/2-\epsilon} \right)$ for some $\epsilon > 0$.  

(E.35)  

Where the last result follows from the bandwidth convergence conditions in Assumption 1.3.9. Next, let

$$U^b_{L_2} = \frac{1}{L_2 h^c_{x, L_2}} \sum_{i=1}^{L_2} \sum_{t=1}^{L_2} \int_{t \in \mathcal{G}(X_i)} \varphi^{F_0,k}(P_t, X_t, t, X_i; h_{tb, L_2}, h_{x, L_2}) 1 \left[ \phi(t | X_i) \geq 0 \right] dQ(t | X_i) W_X(X_i).$$

And note (from (A.23)) that $E_{2,(p,X)} \left[ \varphi^{F_0,k}(P_t, X_t, t, X_i; h_{tb, L_2}, h_{x, L_2}) \mid X_i \right] = 0$. Fix $p$ and $x$ and let

$$\Xi^{F_0,k}_{2}(p, x, X^c, h_{tb, L_2}, h_{x, L_2}) =$$

$$E_{2,X^c} \left[ \int_{t \in \mathcal{G}(X)} \varphi^{F_0,k}(p, x, t, X; h_{tb, L_2}, h_{x, L_2}) 1 \left[ \phi(t | X) \geq 0 \right] dQ(t | X) W_X(X) \mid X^c \right],$$

$$\lambda^{F_0,k}_{2}(P_i, X_i; h_{tb, L_2}, h_{x, L_2}) =$$

$$\int_{t \in [-S, S]} \Xi^{F_0,k}_{2}(P_i, X_i, h_{x, L_2} \psi + X^c_i; h_{tb, L_2}, h_{x, L_2}) f_{2,X^c}(h_{x, L_2} \psi + X^c_i) d\psi$$

103
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,

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

\[
E_{2,X} \left[ E_{2,(P,X)} \left[ \int_{t \in H(X_i)} \phi_0 (P, t, X_i; h_{p,L_2}, h_{x,L_2}) \right] X_i \right] \mathbb{1} \left[ \phi_k (t | X_i) \geq 0 \right] dQ(t | X_i) W_{X_i} (X_i) \]

= 0

The Hoeffding decomposition of \( U_{L_2}^b \) yields the following result,

\[
U_{L_2}^b = \frac{1}{L_2} \sum_{i=1}^{L_2} \lambda_2^{F_{0,k}} (P_i, X_i; h_{p,L_2}, h_{x,L_2}) + \tilde{\nu}^b,
\]

where \( \tilde{\nu}^b = O_p \left( \frac{1}{L_2 h_{x,L_2}^c} \right) = O_p \left( L_2^{-1/2-\epsilon} \right) \) for some \( \epsilon > 0 \).

(A.36)

Once again the last line results from the bandwidth properties described in Assumption 1.3.9. Next let us analyze the generalized U-statistic\(^1\) (see Section 5.1.3 in Serfling (1980)) ,

\[
U_{L_1,L_2}^c = \frac{1}{L_2 L_1 h_{x,L_2}^c} \sum_{i=1}^{L_2} \sum_{m=1}^{L_1} \int_{t \in H(X_i)} \phi_1 (b_m, X_m, t, X_i; h_{b,L_1}, h_{x,L_1}) \mathbb{1} \left[ \phi_k (t | X_i) \geq 0 \right] dQ(t | X_i) W_{X_i} (X_i).
\]

\(^1\)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).
Note that $E_1(b, X) \left[ \varphi^{G_1}(b_m, X_m, t, X_i; h_{b,L_1}, h_{x,L_1}) | X_i \right] = 0$ (again, see the paragraph preceding Equation (A.1)). Fix $b$ and $x$ and let

$$
\Xi^{G_1}_2(b, x, X^c, h_{b,L_1}, h_{x,L_1}) = 
\begin{array}{c}
E_{2,X^c}[ \int_{t \in \mathcal{F}(X)} \varphi^{G_1}(b, x, t, X; h_{b,L_1}, h_{x,L_1}) 1 [\phi_k(t|X) \geq 0] dQ(t|X) W_{x}(X) | X^c],
\end{array}
$$

$$
\lambda^{G_1}_2(b, X; h_{b,L_1}, h_{x,L_1}) = 
\begin{array}{c}
\Xi^{G_1}_2(b, X, h_{b,L_1}, h_{x,L_1})
\end{array}
\int_{\psi \in [-S,S]^c} \Xi^{G_1}_2(b, X, h_{b,L_1}, h_{x,L_1}) f_{2,X^c}(h_{b,L_1}, h_{x,L_1}) d\psi
$$

As in the previous cases, the last functional satisfies $E_1(b, X) \left[ \lambda^{G_1}_2(b, X; h_{b,L_1}, h_{x,L_1}) \right] = 0$. This can be shown by iterated expectations, since

$$
\begin{array}{c}
E_1(b, X) \left[ \lambda^{G_1}_2(b, X_i; h_{b,L_1}, h_{x,L_1}) \right] = 
E_2, X \left[ E_1(b, X) \left[ \int_{t \in \mathcal{F}(X)} \varphi^{G_1}(b, X, t, X_i; h_{b,L_1}, h_{x,L_1}) 1 [\phi_k(t|X_i) \geq 0] dQ(t|X_i) W_{x}(X_i) \right] = 0
\end{array}
$$

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

$$
U^c_{L_1,L_2} = \frac{1}{L_1} \sum_{i=1}^{L_1} \lambda^{G_1}_2(b_i, X_i; h_{b,L_1}, h_{x,L_1}) + \tilde{\nu}^c,
$$

where $\tilde{\nu}^c = O_p \left( \frac{1}{\sqrt{L_1 L_2 h^c_{x,L_1}}} \right) = \left\{ \begin{array}{ll}
O_p \left( L_1^{-1/2-\epsilon} \right) & \text{for some } \epsilon > 0.
\end{array} \right.$

(A.37)

The last result being a consequence of Assumption 1.3.9. The last term left to simplify the asymptotic properties of expressions (A.32) and (A.33) is the following generalized U-statistic,

$$
U^d_{L_1,L_2} = 
\frac{1}{L_1 L_2 h^c_{x,L_2}} \sum_{i=1}^{L_1} \sum_{\ell=1}^{L_2} \int_{t \in \mathcal{F}(X_i)} \psi^{F_0,k}(P_\ell, X_\ell, t, X_i; h_{p,L_2}, h_{x,L_2}) 1 [\phi_k(t|X_i) \geq 0] dQ(t|X_i) \cdot W_{x}(X_i)
$$

105
As it was the case with the statistics $U^a_{L1} \cdot U^b_{L2}$ and $U^c_{L1,L2}$, from Equation (A.23) we have in this case that $E_{2,X} \cdot \left[ \phi_{F_0,k} (P_i, X_i, t, X_i; h_{p,L2}, h_{x,L2}) \right] = 0$. Fix $p, x$ and let

$$
\Xi_{F_0,k}^L (P, X, h_{p,L2}, h_{x,L2}) = 
\int_{t \in \mathcal{B}(X)} \phi_{F_0,k} (P, x, t, X; h_{p,L2}, h_{x,L2}) \left[ \Xi_{F_0,k}^L (P, X, h_{p,L2}, h_{x,L2}) \right] = 0.
$$

The last functional satisfies $E_{2,(P,X)} \left[ \lambda_{F_0,k}^L (P, X, h_{p,L2}, h_{x,L2}) \right] = 0$ since, by iterated expectations,

$$
E_{2,(P,X)} \left[ \lambda_{F_0,k}^L (P, X, h_{p,L2}, h_{x,L2}) \right] = 
E_{1,X} \left[ \int_{t \in \mathcal{B}(X)} E_{2,(P,X)} \left[ \phi_{F_0,k} (P, X, t, X; h_{p,L2}, h_{x,L2}) \right] \right] = 0.
$$

The Hoeffding decomposition expression for $U^d_{L1,L2}$ is,

$$
U^d_{L1,L2} = \frac{1}{L_2} \sum_{i=1}^{L_2} \lambda_{F_0,k}^L (P, X, h_{p,L2}, h_{x,L2}) + \hat{\nu}^d,
$$

where $\hat{\nu}^d = O_p \left( \frac{1}{\sqrt{L_1 L_2 h_{x,L2}^2}} \right) = \begin{cases} 
O_p \left( L_1^{-1/2-\epsilon} \right) & \text{for some } \epsilon > 0 \\
O_p \left( L_2^{-1/2-\epsilon} \right) & \text{for some } \epsilon > 0
\end{cases}$. 

(A.38)
Combining the previous results with \((A.33)\),

\[
\frac{1}{L_1} \sum_{i=1}^{L_1} \tilde{\Lambda}_{k,1}(X_i) W_{X}(X_i) = E_{1,X} [\Lambda_k(X_i) W_{X}(X_i)] \\
+ \frac{1}{L_1} \sum_{i=1}^{L_1} (\Lambda_k(X_i) W_{X}(X_i) - E_{1,X} [\Lambda_k(X_i) W_{X}(X_i)]) \\
+ \frac{1}{L_2} \sum_{i=1}^{L_2} \lambda_{1}^{F_0,k}(P_i, X_i; h_{p,L_2}, h_{x,L_2}) - \frac{1}{L_1} \sum_{i=1}^{L_1} \lambda_{1}^{G_1}(b_i, X_i; h_{b,L_1}, h_{x,L_1}) + \tilde{\nu}_1,
\]

where \(\tilde{\nu}_1 = O_p \left( L_1^{-1/2-\epsilon} \right) + O_p \left( L_1^{-1/2-\epsilon} \right) \) for some \(\epsilon > 0\).

And Equation \((A.34)\) becomes,

\[
\frac{1}{L_2} \sum_{i=1}^{L_2} \tilde{\Lambda}_{k,2}(X_i) W_{X}(X_i) = E_{2,X} [\Lambda_k(X_i) W_{X}(X_i)] \\
+ \frac{1}{L_2} \sum_{i=1}^{L_2} (\Lambda_k(X_i) W_{X}(X_i) - E_{2,X} [\Lambda_k(X_i) W_{X}(X_i)]) \\
+ \frac{1}{L_2} \sum_{i=1}^{L_2} \lambda_{2}^{F_0,k}(P_i, X_i; h_{p,L_2}, h_{x,L_2}) - \frac{1}{L_1} \sum_{i=1}^{L_1} \lambda_{2}^{G_1}(b_i, X_i; h_{b,L_1}, h_{x,L_1}) + \tilde{\nu}_2,
\]

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

Let

\[
\psi_{1,k}(b_j, X_j; h_{b,L_1}, h_{x,L_1}) = (\Lambda_k(X_j) W_{X}(X_j) - E_{1,X} [\Lambda_k(X_j) W_{X}(X_j)]) \\
- \lambda_{1}^{G_1}(b_j, X_j; h_{b,L_1}, h_{x,L_1}) - \lambda_{2}^{G_1}(b_j, X_j; h_{b,L_1}, h_{x,L_1}),
\]

\[
\psi_{2,k}(P_i, X_i; h_{p,L_2}, h_{x,L_2}) = (\Lambda_k(X_i) W_{X}(X_i) - E [\Lambda_k(X_i) W_{X}(X_i)]) \\
+ \lambda_{1}^{F_0,k}(P_i, X_i; h_{p,L_2}, h_{x,L_2}) + \lambda_{2}^{F_0,k}(P_i, X_i; h_{p,L_2}, h_{x,L_2}).
\]
Let $L \equiv L_1 + L_2$. Then, for some $\Delta > \frac{1}{2}$,

\[
\hat{T}_k = T_k + \frac{1}{L_1} \sum_{j=1}^{L_1} \psi_{1,k}(b_j, X_j; h_{b,L_1}, h_{x,L_1}) + \frac{1}{L_2} \sum_{i=1}^{L_2} \psi_{2,k}(P_i, X_i; h_{p,L_2}, h_{x,L_2}) + o_p(L_1^{-\Delta}) + o_p(L_2^{-\Delta})
\]

This proves Theorem 1.3.1.
Appendix B
Exploring evidence of collusion in the sample of first-price auctions: a reduced-form analysis

As it was mentioned in Section 1.4.6, recent econometric results show that the distribution of values in ascending auctions can still be identified and estimated nonparametrically in the presence of collusion. However, collusion in the sample of first-price auctions would invalidate the bounds and the tests for $k$-rationalizability. For this reason it is useful to perform at least an exploratory analysis for the presence of collusion in the sample of first-price auctions. There has been extensive work on the presence of collusion in auctions (Graham and Marshall (1987), Porter and Zona (1993), Baldwin et al. (1997), Marshall and Marx (2007), Marshall and Marx (2009), Baldwin et al. (1997), Athey et al. (2011), Kaplan et al. (2017), Schurter (2017a)). Specifically, the possibility of collusion in timber auctions has been explored before (Baldwin et al. (1997), Athey et al. (2011), Schurter (2017a)), but not in the specific context of the time period and regions analyzed in my data set. While the focus of my paper is not collusion, but to develop formal tests for rationalizability, as part of my analysis I perform a reduced-form study similar to the one undertaken by Porter and Zona (1993) (henceforth PZ93) who studied auctions for highway construction contracts. To facilitate the illustration, I use slightly different notation in this section relative to the rest of my paper. Following PZ93, I use the following approximation
for bidding behavior in first-price auctions,

$$\log(b_{it}) = \alpha'X_t + \beta'X_{it} + \epsilon_{it},$$  \hspace{1cm} (B.1)

where $X_t$ denotes observable characteristics at the auction level and $X_{it}$ denotes observable characteristics of bidder $i$ at the time of auction $t$. The term $\epsilon_{it}$ captures private information observed by bidder $i$ in auction $t$.

**B.1 Analysis based on bid levels**

Let the superscript ‘$c$’ indicate a potential colluder and $nc$ indicate a bidder assumed to be bidding competitively. I can generalize the bidding rule described above as

$$\log(b_{it}) = \alpha^cX_t + \beta^cX_{it} + \epsilon_{it} \text{ for (potential) colluders},$$

$$\log(b_{it}) = \alpha^{nc}X_t + \beta^{nc}X_{it} + \epsilon_{it} \text{ for competitive bidders}. \hspace{1cm} (B.2)$$

Grouping $\theta \equiv (\alpha, \beta)$, the true presence of colluding behavior among potential colluders would require $\theta^c \neq \theta^{nc}$. Therefore, an exploratory “test” for the presence of collusion would be a test of the null hypothesis $H_0 : \theta^c = \theta^{nc}$. Obviously, implementing this test requires having a candidate set of colluders. For the time period and regions I analyze, I do not have a comprehensive legal study or precedents that can allow me to definitively identify a specific set of potential colluders, so I do an exploratory analysis instead. I identify candidate sets of colluders by focusing on bidders who have won a larger-than-usual number of “high-value” auctions. I proceed in the following steps:

1. I define a “high-value” (first price) auction as any auction where transaction price (in real 2010 USD) was higher than the median value of all bids submitted during the time period analyzed. Here, value and bids are measured in total volume, not per-volume.

2. I look at a (super)set of potential colluders as the collection of bidders in the top $D^{th}$ quantile in terms of number of high-valued auctions won during the time period of my sample.

3. For a given $D$, I estimate (B.2) separately for the subsample of potential colluders and the subsample of competitive bidders using OLS. Following PZ93,
I focus on those auctions where at least two competitive bidders participated. I then test $H_0: \theta^e = \theta^{nc}$ using a Wald test with a robust estimator of the asymptotic variance of my estimators.

4. I repeat the exercise for different values of $D$

I include similar regressors in $X_t$ and $X_{it}$ to those used in Chapter 1 of Shen (2017), who performed this analysis for timber auctions in a different time period and region than my sample.

B.1.1 Variables included

The variables included in $X_t$ and $X_{it}$ are very similar to those used in Chapter 1 of Shen (2017), who performed this analysis for timber auctions in a different time period and region than my sample. The auction-specific regressors $X_t$ I include are the following (all monetary figures are measured in real 2010 USD):

- Number of bidders: Those who submitted bids at least equal to the reserve price in the auction.
- Total volume in mbf (thousands of board feet), in logarithm.
- Duration of contract: Maximum number of months the winner can take to clear-cut the tract, normalized by volume.
- Road construction cost (normalized by volume).
- Forest inventory: The difference between the total volume that has been previously auctioned in the forest – during the time period of my data –, minus the estimate of the total amount harvested.
- Reserve price (normalized by volume).
- Density of timber in the tract: Measured in mbf per acre, in logarithm.
- An indicator for Willamette National Forest (in Oregon), which presented the highest concentration of high-value auctions.

The bidder-auction observable characteristics $X_{it}$ include the following:
• Business size: Split into 8 categories according to number of employees. Large firms are those with 500 or more.

• Bidder inventory: The estimated amount of timber harvested by bidder \( i \), at the time of auction \( i \) from all previous auctions won by \( i \).

• The number of previous auctions won by bidder \( i \).

• The total number of previous auctions where bidder \( i \) participated.

To compute the amount of timber harvested at time period \( t \) from auction \( \ell \) (whose contract was awarded prior to \( t \)), I follow the approach in Chapter 1 of Shen (2017) (which, in turn, follows the standard approach in timber auctions). Let \( t_\ell \) and \( \bar{t}_\ell \) be the starting and ending periods of the auction-contract. I assume that harvest takes place during the last \( 1/3 \) of the duration of the contract, so that the starting date for harvest is \( t_0 = t_\ell + (1/3) \cdot (\bar{t}_\ell - t_\ell) \), then I assume a uniform rate of harvest. Therefore, if \( VOL_\ell \) is the total volume of auction \( \ell \) whose contract runs from \( t_\ell \) to \( \bar{t}_\ell \), then the amount harvested at time period \( t \) is

\[
HARVEST_{\ell t} = 1 \left[ t_0 \leq t \leq \bar{t}_\ell \right] \cdot \left( \frac{t - t_0}{\bar{t}_\ell - t_0} \right) \cdot VOL_\ell + 1 \left[ t > \bar{t}_\ell \right] \cdot VOL_\ell.
\]

Let \( A_{i,t} \) be the collection of auctions won by bidder \( i \) prior to time period \( t \) and \( A_{f,t} \) be the collection of auctions involving forest \( f \) prior to time period \( t \). Then, the total inventory of bidder \( i \) at time period \( t \) is \( INV_{i,t} = \sum_{\ell \in A_{i,t}} HARVEST_{\ell t} \), and the total amount harvested in forest \( f \) at time period \( t \) is \( HARV_{f,t} = \sum_{\ell \in A_{f,t}} HARVEST_{\ell t} \).

Let \( A_{i,t} \) be the collection of auctions won by bidder \( i \) prior to time period \( t \) and \( A_{f,t} \) be the collection of auctions involving forest \( f \) prior to time period \( t \). Then, the total inventory of bidder \( i \) at time period \( t \) is \( INV_{i,t} = \sum_{\ell \in A_{i,t}} HARVEST_{\ell t} \), and the total amount harvested in forest \( f \) at time period \( t \) is \( HARV_{f,t} = \sum_{\ell \in A_{f,t}} HARVEST_{\ell t} \).

The inventory available for harvest in forest \( f \) at time period \( t \) would be \( INV_{f,t} = \sum_{\ell \in A_{f,t}} VOL_\ell - HARV_{f,t} \).

The regressors included in my analysis are those described above, plus a constant. Thus, \( \theta \) includes 13 parameters in total.

Table B.1 presents the results of the Wald test for \( H_0 : \theta^c = \theta^{nc} \) for different values of \( D \). Collusive behavior would imply a rejection of \( H_0 \).

Looking at the results from Table B.1, while I could reject \( H_0 : \theta^c = \theta^{nc} \) at a 5% significance level\(^1\), I fail to reject it at a 1% level. By its limited nature, the above

\(^1\)Notice that, by construction, the candidate sets for colluders are nested inside the largest \( D \), so the joint null hypothesis that there is no collusion for \( D = 1, 5, 10, 15 \) and 20 is simply the hypothesis that there is no collusion for \( D = 20 \). This is why the p-values shown in the table are not Bonferroni-corrected, which would lead to even larger p-values.
Table B.1. Collusion analysis based on bid levels: Results from the test $H_0 : \theta^c = \theta^{nc}$ where the set of potential colluders is defined as those bidders in the top $D^{th}$ quantile in terms of high-valued auctions won during the time period.

<table>
<thead>
<tr>
<th>$D$</th>
<th>Wald statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>13.5018</td>
<td>0.4098</td>
</tr>
<tr>
<td>5%</td>
<td>18.3468</td>
<td>0.5774</td>
</tr>
<tr>
<td>10%</td>
<td>27.0183</td>
<td>0.0124</td>
</tr>
<tr>
<td>15%</td>
<td>22.2216</td>
<td>0.0520</td>
</tr>
<tr>
<td>20%</td>
<td>18.3468</td>
<td>0.1447</td>
</tr>
</tbody>
</table>

The analysis is not in any way sufficient to rule out the presence of collusion, but the spirit of the analysis was to cast a wide enough net that, if the presence of collusive behavior is significant enough, it should be reflected in the results of the tests. For example, in contrast to the results shown above, the p-value for the equivalent test in Chapter 1 of Shen (2017) was $5.8 \times 10^{-8}$, although it should be pointed out that in that specific data set there was prior information about the potential colluders.

B.2 Analysis based on bid rankings

In order to detect the presence of phantom bidding, PZ93 also suggest tests based on the rankings of bids (in addition to the analysis on levels). The idea is that in the presence of collusion, fundamental differences should arise between the ordering of competitive and cartel bids. Here I go back to the approximation in (B.1). If I now assume that the private information shocks $\epsilon_{it}$ have an i.i.d Type-I extreme value distribution, then the likelihood of observing a particular ranking of bids $b_{r_{1t}} > b_{r_{2t}} > \cdots > b_{rn_{1t}}$ in auction $t$, conditional on observable characteristics, is

$$P(b_{r_{1t}} > b_{r_{2t}} > \cdots > b_{rn_{1t}}) = \prod_{i=1}^{n_t} \frac{e^{\theta'X_{it}}}{\sum_{j=1}^{n_t} e^{\theta'X_{jt}}}.$$

Notice that the term $\alpha'X_t$, which captures the observed heterogeneity at the auction level, drops out. This constitutes the type of rank-ordered logit model analyzed in Hausman and Ruud (1987). If the model is properly approximated by (B.1), estimators based on different subsets of ranks (for example, the ranks of the two highest bids only, or the ranks of the $n_t - 2$ lowest ranks only) should converge
asymptotically to the true parameter value for $\beta$ (see Equation 5 in Hausman and Ruud (1987)). In particular, under the null hypothesis of no phantom bidding, parameters estimated using only the lowest ranks of bids submitted by potential colluders should be asymptotically equal to those estimated using the highest ranks. To perform an exploratory analysis of phantom bids using rank-based tests I proceeded in the following steps:

1. I identify tentatively a (super)set of potential colluders following the same approach as before, by looking at the top $D^{th}$ quantile in terms of number of high-valued auctions won during the time period of my sample.

2. I estimate the parameters of the rank-ordered logit model using only the ranks of the top-two bids submitted by the set of potential colluders.

3. I repeat the estimation using the ranks of the $n_t - 3$ lowest bids submitted by the set of potential colluders.

4. I perform a Wald test for the null hypothesis that both sets of parameters are equal. Under the presence of phantom bidding, this null hypothesis should be rejected.

5. I repeat the exercise for the set of non-colluders as a way to check for specification error (if the null hypothesis is rejected, it would be evidence against the specification of the model).

Implementation of the test requires me to focus on auctions where at least three potential colluders submitted a bid, resulting in relatively small sample sizes. In particular, 3 or more potential colluders submitted bids in less than 60 auctions in my sample when I use $D < 10\%$. The results in Table B.2 appear consistent with our analysis on bid-levels: At a 5\% significance level, evidence of noncompetitive bidding appears more compelling when I focus on bidders in the upper 10\% percentile of high-value auction wins. However, bidding behavior appears consistent with competition if I use a 1\% significance level.

B.3 Summary of reduced-form collusion analysis

For the (super)sets of potential colluders I have analyzed here:
Table B.2. Collusion analysis based on bid rankings: Testing $H_0$ : high-bid and low-bid rankings are generated by the same bidding rule. Set of potential colluders is defined as those bidders in the top $D^{th}$ quantile in terms of high-valued auctions won during the time period.

<table>
<thead>
<tr>
<th></th>
<th>Potential Colluders</th>
<th>Non-Colluders</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wald statistic</td>
<td>p-value</td>
</tr>
<tr>
<td>$D = 10%$</td>
<td>14.655</td>
<td>0.012</td>
</tr>
<tr>
<td>$D = 15%$</td>
<td>6.594</td>
<td>0.252</td>
</tr>
<tr>
<td>$D = 20%$</td>
<td>6.686</td>
<td>0.245</td>
</tr>
</tbody>
</table>

Note: Using $D < 10\%$ resulted in fewer than 60 first-price auctions with 3 or more potential colluders in my sample. Therefore the test was implemented only for $D \geq 10\%$.

(i) At a 5\% significance level, evidence consistent with noncompetitive bidding appears when we focus on bidders in the upper 10\% percentile of high-value auction wins.

(ii) Bidding behavior appears consistent with competition in all cases analyzed if I use a 1\% significance level.

(iii) Other studies that have found evidence of collusion using these types of tests (Shen (2017), Porter and Zona (1993)) have found significantly stronger evidence, being able to reject competition with p-values close to zero.

(iv) Under the maintained assumption of symmetry, the deviations from Nash Equilibrium bidding that I uncover in my structural analysis cannot be explained by collusion, since bids appear to be above their Nash Equilibrium levels with positive probability.

It must be reiterated that, like the analysis in PZ93, these findings depend crucially on the approximation used for bids (Equation (B.1)), where an effort was made to include a rich collection of auction-specific and bidder-auction specific covariates. Furthermore, having precise empirical-based (e.g, based on legal findings) information about the identities of cartel firms in my specific region and time period studied could lead to different results. This should be understood only as an exploratory analysis.
Appendix C
Omitted proofs from chapter 3

C.1 Proof of theorem (3.3.1)

To characterize the limiting properties of the estimator \( \hat{R}(t) \), consider:

\[
\hat{R}(t) - R(t) = \left[ \frac{1}{nh_n} \sum_{i=1}^{n} Y_i K \left( \frac{X_i \hat{\beta} - t}{h_n} \right) - \frac{1}{nh_n} \sum_{i=1}^{n} Y_i K \left( \frac{X_i \beta^* - t}{h_n} \right) \right] - \left[ \frac{1}{nh_n} \sum_{i=1}^{n} Y_i K \left( \frac{X_i \beta^* - t}{h_n} \right) - \mathbb{E}[Y|T = t] f_T(t) \right]
\]

(C.1)

Consider the first term:

\[
I = \frac{1}{nh_n} \sum_{i=1}^{n} Y_i K \left( \frac{X_i \hat{\beta} - t}{h_n} \right) - \frac{1}{nh_n} \sum_{i=1}^{n} Y_i K \left( \frac{X_i \beta^* - t}{h_n} \right) =
\]

\[
= \frac{1}{nh_n} \sum_{i=1}^{n} Y_i \left( K \left( \frac{X_i \hat{\beta} - t}{h_n} \right) - K \left( \frac{X_i \beta^* - t}{h_n} \right) \right)
\]

Taylor expansion of \( K \left( \frac{X_i \hat{\beta} - t}{h_n} \right) \) around \( \hat{\beta} \) is

\[
K \left( \frac{X_i \hat{\beta} - t}{h_n} \right) = K \left( \frac{X_i \beta^* - t}{h_n} \right) + K' \left( \frac{X_i \beta^* - t}{h_n} \right) \frac{X_i}{h_n} (\hat{\beta} - \beta^*) +
\]
\[
+ \frac{1}{2} (\beta - \beta^*)' K'' \left( X_i' \beta - t \right) \frac{X_i X_i'}{h_n^2} (\beta - \beta^*),
\]

where \( \beta \) is a random vector between \( \hat{\beta} \) and \( \beta^* \).

Since \( \beta - \beta^* = \frac{1}{n} \sum_{j=1}^n \psi_j + o_p(n^{-\frac{1}{2}}) \) and \( \psi_j \) are influence functions, \( \beta - \beta^* = O_p(n^{-\frac{1}{2}}) \). Thus, the first term in (C.1) has the following form:

\[
I = \frac{1}{n h_n} \sum_{i=1}^n Y_i K' \left( \frac{X_i' \beta^* - t}{h_n} \right) \frac{X_i' 1}{h_n} n \sum_{j=1}^n \psi_j^\beta +
\]

\[
+ \frac{1}{n h_n} \sum_{i=1}^n Y_i K' \left( \frac{X_i' \beta^* - t}{h_n} \right) \frac{X_i' h_n}{h_n^2} \sigma_o(n^{-\delta}) + \frac{1}{n h_n} \sum_{i=1}^n Y_i \frac{1}{2} K'' \left( \frac{X_i' \beta - t}{h_n} \right) \frac{X_i X_i'}{h_n^2} O_p(n^{-1}) =
\]

\[
O_p(1)
\]

\[
O_p(n^{-\delta - \frac{1}{2} h_n^{-2}})
\]

\[
O_p(n^{-\frac{3}{2} h_n^{-3}})
\]

\[
= \frac{1}{n^2 h_n} \sum_{i=1}^n \sum_{j=1}^n Y_i K' \left( \frac{X_i' \beta^* - t}{h_n} \right) \frac{X_i' \psi_j^\beta + o_p(n^{-\delta - \frac{1}{2} h_n^{-2}}) + O_p(n^{-\frac{3}{2} h_n^{-3}}).}
\]

Consider the second term in (C.1)

\[
II = \frac{1}{n h_n} \sum_{i=1}^n Y_i K \left( \frac{X_i' \beta^* - t}{h_n} \right) - \mathbb{E} [Y|T = t] f_T(t) =
\]

\[
= \frac{1}{n} \sum_{i=1}^n \left( \frac{1}{h_n} Y_i K \left( \frac{X_i' \beta^* - t}{h_n} \right) - \mathbb{E} \left[ \frac{1}{h_n} Y K \left( \frac{X' \beta^* - t}{h_n} \right) \right] \right) +
\]

\[
+ \mathbb{E} \left[ \frac{1}{h_n} Y K \left( \frac{X' \beta^* - t}{h_n} \right) \right] - \mathbb{E} [Y|T = t] f_T(t)
\]

The first expectation in \( II_b \):

\[
\mathbb{E} \left[ \frac{1}{h_n} Y K \left( \frac{X' \beta^* - t}{h_n} \right) \right] = \mathbb{E} \left[ \frac{1}{h_n} Y K \left( \frac{T - t}{h_n} \right) \right] = \left[ \frac{1}{h_n} \mathbb{E} [Y|T] K \left( \frac{T - t}{h_n} \right) \right] =
\]

\[
= \int \frac{1}{h_n} \mathbb{E} [Y|T = u] K \left( \frac{u - t}{h_n} \right) f_T(u) du = \int \frac{1}{h_n} \mu(u) K \left( \frac{u - t}{h_n} \right) f_T(u) du =
\]

\[
= \int \mu(t + h_n v) K (v) f_T(t + h_n v) dv
\]
Taylor expansion of $\mu(t + h_n v) f_T(t + h_n v)$ around $t$:

$$
\mu(t + h_n v) f_T(t + h_n v) = \mu(t) f_T(t) + \left( \mu(t + \tilde{h}_n v) f_T(t + \tilde{h}_n v) \right)' h_n v,
$$

where $\tilde{h}_n$ between 0 and $h_n$. Thus, the first expectation is

$$
\mathbb{E} \left[ \frac{1}{h_n} YK \left( \frac{X' \beta^* - t}{h_n} \right) \right] = \mu(t) f_T(t) \int K(v) dv + h_n \int \left( \mu(t + \tilde{h}_n v) f_T(t + \tilde{h}_n v) \right)' K(v) dv =
$$

$$
= \mu(t) f_T(t) + O(h_n^M)
$$

Bias-reducing kernel of order $M$, s.t. $\int K(v) dv = 1$, $\int v^j K(v) dv = 0$, for all $j = 1, \ldots, M - 1$ and $\int |v|^M K(v) dv < \infty$

Thus, $IIb = O(h_n^M)$.

Consider term $IIa$. It is a sum of random variables with zero mean and variance:

$$
\mathbb{V} \left[ \frac{1}{h_n} YK \left( \frac{X' \beta^* - t}{h_n} \right) \right] = \frac{1}{h_n^2} \mathbb{V} \left[ YK \left( \frac{X' \beta^* - t}{h_n} \right) \right] =
$$

$$
= \frac{1}{h_n^2} \left( \mathbb{E} \left[ Y^2 K^2 \left( \frac{X' \beta^* - t}{h_n} \right) \right] - \mathbb{E}^2 \left[ YK \left( \frac{X' \beta^* - t}{h_n} \right) \right] \right) =
$$

$$
= \frac{1}{h_n^2} \left( \mathbb{E} \left[ Y^2 |T| \right] K^2 \left( \frac{T - t}{h_n} \right) \right) - \mathbb{E}^2 \left[ YK \left( \frac{T - t}{h_n} \right) \right] =
$$

$$
= \frac{1}{h_n^2} \left( \mathbb{E} \left[ Y^2 |T| \right] K^2 \left( \frac{T - t}{h_n} \right) \right) - \mathbb{E}^2 \left[ YK \left( \frac{T - t}{h_n} \right) \right].
$$

Denote $\sigma_{Y|T} = \mathbb{E} [Y^2 |T]$.

$$
\mathbb{V} \left[ \frac{1}{h_n} YK \left( \frac{X' \beta^* - t}{h_n} \right) \right] = \frac{1}{h_n^2} \left( \int \sigma(u) K^2 \left( \frac{u - t}{h_n} \right) f_T(u) du - (h_n \mu(t) f_T(t) + O(h_n^{M+1}))^2 \right) =
$$

$$
= \frac{1}{h_n} \int \sigma(t + h_n v) K^2(v) f_T(t + h_n v) dv + O(1)
$$

Taylor expansion of $\sigma(t + h_n v) f_T(t + h_n v)$ around $t$:

$$
\sigma(t + h_n v) f_T(t + h_n v) = \sigma(t) f_T(t) + \left( \sigma(t + \tilde{h}_n v) f_T(t + \tilde{h}_n v) \right)' h_n v,
$$

where $\tilde{h}_n$ between 0 and $h_n$. 

118
\[
\mathbb{V} \left[ \frac{1}{h_n} Y K \left( \frac{X' \beta^* - t}{h_n} \right) \right] = \\
\frac{1}{h_n} \left( \sigma(t) f_T(t) \int K^2(v) dv + h_n \int (\sigma(t + \tilde{h}_n v) f_T(t + \tilde{h}_n v))' K^2(v) dv \right) + O(1) = \\
\frac{1}{h_n} \sigma(t) f_T(t) \int K^2(v) dv + O(1).
\]

Therefore,

\[
II = \frac{1}{nh_n} \sum_{i=1}^{n} \left( Y_i K \left( \frac{X'_i \beta^* - t}{h_n} \right) - E \left[ Y K \left( \frac{X' \beta^* - t}{h_n} \right) \right] \right) + O(h_n^M).
\]

Thus,

\[
\tilde{R}(t) - R(t) = \frac{1}{n^2 h_n} \sum_{i=1}^{n} \sum_{j=1}^{n} Y_i K' \left( \frac{X'_i \beta^* - t}{h_n} \right) \frac{X'_i}{h_n} \psi_j^\beta - \\
- \frac{1}{nh_n} \sum_{i=1}^{n} \left( Y_i K \left( \frac{X'_i \beta^* - t}{h_n} \right) - E \left[ Y K \left( \frac{X' \beta^* - t}{h_n} \right) \right] \right) + \\
+ O(h_n^M) + o_p(n^{-1} h_n^2) + O_p(n^{-1} h_n^{-3})
\] (C.2)

Denote \( W_i = (Y_i, X_i) \). Consider double sum in (C.2)

\[
\frac{1}{n^2 h_n} \sum_{i=1}^{n} \sum_{j=1}^{n} Y_i K' \left( \frac{X'_i \beta^* - t}{h_n} \right) \frac{X'_i}{h_n} \psi_j^\beta \equiv \\
\frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} \frac{1}{n} \left( Y_i K' \left( \frac{X'_i \beta^* - t}{h_n} \right) \frac{X'_i}{h_n^2} \psi_j^\beta + Y_j K' \left( \frac{X'_j \beta^* - t}{h_n} \right) \frac{X'_j}{h_n^2} \psi_i^\beta \right) + \\
+ \frac{1}{n^2} \sum_{i=1}^{n} Y_i K' \left( \frac{X'_i \beta^* - t}{h_n} \right) \frac{X'_i}{h_n^2} \psi_i^\beta
\]

\[119\]
The second term in the double sum is

$$\frac{1}{n^2} \sum_{i=1}^{n} Y_i K' \left( \frac{X_i' \beta^* - t}{h_n} \right) X_i' \psi_i = \frac{1}{nh_n^2} \sum_{i=1}^{n} Y_i K' \left( \frac{X_i' \beta^* - t}{h_n} \right) X_i' \psi_i$$

1. Assume that $K'(u) \leq K$ and $E[|YX'\psi\beta|] < \infty$ then the second term is $O_p(n^{-\frac{3}{2}}h_n^{-2})$.

Denote $g(W_i, W_j) = \frac{1}{2} \left(Y_i K' \left( \frac{X_i' \beta^* - t}{h_n} \right) X_i' \psi_i + Y_j K' \left( \frac{X_j' \beta^* - t}{h_n} \right) X_j' \psi_j \right)$ and $E[g(W_i, W_j)] = 0$ because $E[\psi\beta] = 0$. Then, the double sum is a symmetric U-statistics. Using Hoeffding decomposition of this U-statistics, its leading term is:

$$E[g(W_i, W_j) | W_i] = \frac{1}{2} \left( \frac{1 - 1}{h_n^2} E \left[ Y K' \left( \frac{X' \beta^* - t}{h_n} \right) X' \right] \psi_i \right)$$

Analyze

$$E \left[ Y K' \left( \frac{X' \beta^* - t}{h_n} \right) X' \right] = E \left[ E[YX' | X' \beta^*] K' \left( \frac{X' \beta^* - t}{h_n} \right) \right]$$

Denote $\nu(t) = E[YX' | X' \beta^* = t]$. Then,

$$E \left[ Y K' \left( \frac{X' \beta^* - t}{h_n} \right) X' \right] = \int \nu(u) K' \left( \frac{u - t}{h_n} \right) f_T(u) du = \int \nu(t + h_n v) K'(v) f_T(t + h_n v) h_n dv.$$

If kernel $K(\cdot)$ is symmetric and the support of $T$ is unbounded then

$$E \left[ Y K' \left( \frac{X' \beta^* - t}{h_n} \right) X' \right] = h_n \nu(t) f_T(t) \int_{0}^{\infty} K'(v) dv + h_n^2 (\nu(t) f_T(t))' \int_{-1}^{0} v K'(v) dv + O(h_n^{M+2})$$

Thus,

$$E[g(W_i, W_j) | W_i] = -\frac{1}{2} \left( \frac{1 - 1}{n} \right) (\nu(t) f_T(t))' \psi_i + O(h_n^{M}) \psi_i$$

So, the leading term in Hoeffding decomposition is of order $O_p(n^{-\frac{3}{2}})$. The second term in Hoeffding decomposition has the form:

$$\frac{2}{n(n-1)} \sum_{i<j}^{n} (g(W_i, W_j) - E[g(W_i, W_j | W_i)] - E[g(W_i, W_j | W_j])] =$$
= \frac{2}{n(n-1)} \sum_{i<j}^{n} \left\{ \frac{1}{2} \right\} \left( Y_i K' \left( \frac{X_i' \beta^* - t}{h_n} \right) \frac{X_i'}{h_n^2} \psi_j^\beta + Y_j K' \left( \frac{X_j' \beta^* - t}{h_n} \right) \frac{X_j'}{h_n^2} \psi_i^\beta \right) - \\
\frac{1}{2} \frac{n-1}{n} \frac{1}{h_n^2} \mathbb{E} \left[ Y K' \left( \frac{X' \beta^* - t}{h_n} \right) X' \right] \psi_i^\beta - \frac{1}{2} \frac{n-1}{n} \frac{1}{h_n^2} \mathbb{E} \left[ Y K' \left( \frac{X' \beta^* - t}{h_n} \right) X' \right] \psi_j^\beta \right} = \\
= \frac{2}{n(n-1)} \sum_{i<j}^{n} \left\{ \frac{1}{2} \right\} \left( Y_i K' \left( \frac{X_i' \beta^* - t}{h_n} \right) \frac{X_i'}{h_n^2} \mathbb{E} \left[ Y K' \left( \frac{X' \beta^* - t}{h_n} \right) X' \right] \psi_j^\beta + \\
+ \frac{1}{2} \frac{n-1}{n} \left( Y_j K' \left( \frac{X_j' \beta^* - t}{h_n} \right) \frac{X_j'}{h_n^2} \mathbb{E} \left[ Y K' \left( \frac{X' \beta^* - t}{h_n} \right) X' \right] \psi_i^\beta \right} \right}

Consider

\mathbb{V} \left[ Y K' \left( \frac{X' \beta^* - t}{h_n} \right) \frac{X'}{h_n^2} \right] = \mathbb{E} \left[ Y^2 \left( K' \left( \frac{X' \beta^* - t}{h_n} \right) \right) \frac{X X'}{h_n^4} \right] - \left( \mathbb{E} \left[ Y K' \left( \frac{X' \beta^* - t}{h_n} \right) \frac{X'}{h_n^2} \right] \right)^2

The first term is

\mathbb{E} \left[ Y^2 \left( K' \left( \frac{X' \beta^* - t}{h_n} \right) \right) \frac{X X'}{h_n^4} \right] = \mathbb{E} \left[ \mathbb{E} \left[ Y^2 XX' | X' \beta^* \right] \frac{1}{h_n^4} \left( \left( \frac{X' \beta^* - t}{h_n} \right) \right)^2 \right]

Denote \eta(t) = \mathbb{E} \left[ Y^2 XX' | X' \beta^* = t \right].

\mathbb{E} \left[ Y^2 \left( K' \left( \frac{X' \beta^* - t}{h_n} \right) \right) \frac{X X'}{h_n^4} \right] = \int \eta(u) \frac{1}{h_n^4} \left( \left( \frac{K' \left( \frac{X' \beta^* - t}{h_n} \right) \right)^2 f_T(u)du = \\
= \int \frac{1}{h_n^3} \eta(t + h_n v) (K'(v))^2 f_T(t + h_n v)dv = \frac{1}{h_n^3} \eta(t) f_T(t) \int (K'(v))^2 dv + \\
+ \frac{1}{h_n^2} (\eta(t)f_T(t))' \int (K'(v))^2 vdv + \frac{1}{h_n} \int (\eta(t + h_n v)f_T(t + h_n v)') v^2 (K'(v))^2 dv = \\
= \frac{1}{h_n^3} \eta(t) f_T(t) \int (K'(v))^2 dv + O \left( \frac{1}{h_n} \right)

Thus,

\mathbb{V} \left[ Y K' \left( \frac{X' \beta^* - t}{h_n} \right) \frac{X'}{h_n^2} \right] = \frac{1}{h_n^3} \eta(t) f_T(t) \int (K'(v))^2 dv + O \left( \frac{1}{h_n} \right) - ((\nu(t)f_T(t))' + O(h_n^M))^2

121
The second term in Hoeffding decomposition is:

\[
\frac{2}{n(n-1)} \sum_{i<j} \left\{ O_p \left( \frac{1}{h_n^3} \right) \psi_i + O_p \left( \frac{1}{h_n^3} \right) \psi_j \right\} = O_p \left( \frac{1}{nh_n^3} \right)
\]

\[
\hat{R}(t) - R(t) = -\frac{1}{n} \sum_{i=1}^{n} \left( \nu(t)f_T(t) \right)' \psi_i + O_p(n^{-\frac{1}{2}}h_n^M) + o_p(n^{-\frac{1}{2}}h_n^M) + O_p(n^{-\frac{3}{2}}h_n^M) - O_p(n^{-\frac{1}{2}})
\]

- \frac{1}{nh_n} \sum_{i=1}^{n} \left( \frac{X_i'\beta^* - t}{h_n} \right) - E \left[ YK \left( \frac{X_i'\beta^* - t}{h_n} \right) \right] + O(h_n^3)

---

**Exact expression:**

\[
\hat{R}(t) - R(t) = -\frac{1}{n} \sum_{i=1}^{n} \left( \nu(t)f_T(t) \right)' \psi_i + O_p(n^{-\frac{1}{2}})
\]

\[
+ \frac{2}{n} \sum_{i=1}^{n} \frac{1}{2} \frac{n-1}{n} \frac{h_n^M}{(M+1)!} \int (\nu(t + \tilde{h}_n)v)f_T(t + \tilde{h}_n)v)^{(M+1)} K'(v) v^{M+1} dv \psi_i + O_p(n^{-\frac{1}{2}}h_n^M)
\]

- first term in Hoeffding decomposition

\[
+ \frac{2}{n(n-1)} \sum_{i<j} \left\{ \frac{1}{2} \frac{n-1}{n} \left( Y_i K' \left( \frac{X_i'\beta^* - t}{h_n} \right) \right) \left( \frac{X_i'}{h_n^2} - E \left[ YK' \left( \frac{X_i'\beta^* - t}{h_n} \right) \frac{X_i'}{h_n^2} \right] \right) \psi_j \right\} + O_p(n^{-1}h_n^{-\frac{3}{2}})
\]

- second term in Hoeffding decomposition

\[
+ \frac{1}{2} \frac{n-1}{n} \left( Y_j K' \left( \frac{X_j'\beta^* - t}{h_n} \right) \right) \left( \frac{X_j'}{h_n^2} - E \left[ YK' \left( \frac{X_j'\beta^* - t}{h_n} \right) \frac{X_j'}{h_n^2} \right] \right) \psi_j + O_p(n^{-1}h_n^{-\frac{3}{2}})
\]

- diagonal elements
\[- \frac{1}{nh_n} \sum_{i=1}^{n} \left( Y_i K \left( \frac{X_i \beta^* - t}{h_n} \right) - \mathbb{E} \left[ Y_i K \left( \frac{X_i \beta^* - t}{h_n} \right) \right] \right) + \mathcal{O}_p \left( \frac{n}{\sqrt{h_n}} \right) \]

\[+ h_n^M \int \left( \mu(t + \tilde{h}_n v) f_T(t + \tilde{h}_n v) \right)' K(v) v^M dv + \mathcal{O}(h_n^M) \]

\[+ \frac{1}{nh_n} \sum_{i=1}^{n} Y_i K' \left( \frac{X_i \beta^* - t}{h_n} \right) \frac{X_i'}{h_n} \mathcal{O}_p \left( n^{-\delta} \right) \]

\[+ \frac{1}{nh_n} \sum_{i=1}^{n} Y_i^2 K'' \left( \frac{X_i \beta^* - t}{h_n} \right) \left[ \frac{X_i'}{h_n^2} \left( \frac{1}{n} \sum_{j=1}^{n} \psi_j^\beta + \mathcal{O}_p \left( n^{-\delta} \right) \right) \right]^2 \]

\[\mathcal{O}_p \left( \frac{n^2}{\sqrt{h_n}} \right) \]

Denote:

\[\psi_i^{R1}(\psi_i^\beta, t) = -n \frac{1}{n} \left( \nu(t) f_T(t) \right)' \psi_i^\beta, \]

\[\psi_i^{R2}(Y_i, X_i, t; h_n) = -Y_i K \left( \frac{X_i \beta^* - t}{h_n} \right) + \mathbb{E} \left[ Y K \left( \frac{X \beta^* - t}{h_n} \right) \right], \]

\[\xi_n^R(t) = \frac{2}{n} \sum_{i=1}^{n} \left\{ \frac{1}{n} \sum_{i<j}^n \left\{ \frac{1}{n} \left( Y_i K' \left( \frac{X_i \beta^* - t}{h_n} \right) \frac{X_i'}{h_n^2} - \mathbb{E} \left[ Y K' \left( \frac{X \beta^* - t}{h_n} \right) \frac{X'}{h_n^2} \right] \right) \psi_j^\beta + \frac{1}{n} \left( Y_j K' \left( \frac{X_j \beta^* - t}{h_n} \right) \frac{X_j'}{h_n^2} - \mathbb{E} \left[ Y K' \left( \frac{X \beta^* - t}{h_n} \right) \frac{X'}{h_n^2} \right] \right) \psi_i^\beta \right\} + \frac{1}{n} \sum_{i=1}^{n} \left( Y_i K' \left( \frac{X_i \beta^* - t}{h_n} \right) X_i \psi_i^\beta + h_n^M \int \left( \mu(t + \tilde{h}_n v) f_T(t + \tilde{h}_n v) \right)' K(v) v^M dv + \frac{1}{nh_n} \sum_{i=1}^{n} Y_i K' \left( \frac{X_i \beta^* - t}{h_n} \right) \frac{X_i'}{h_n} \mathcal{O}_p \left( n^{-\delta} \right) + \frac{1}{nh_n} \sum_{i=1}^{n} Y_i \frac{1}{2} K'' \left( \frac{X_i \beta^* - t}{h_n} \right) \left[ \frac{X_i'}{h_n^2} \left( \frac{1}{n} \sum_{j=1}^{n} \psi_j^\beta + \mathcal{O}_p \left( n^{-\delta} \right) \right) \right]^2 \right\} \]
For the next step in analysis of (3.8), I need:

$$\sup_{t \in \mathcal{T}} |\xi_n^R(t)| = O_p(n^{-\frac{1}{2} - \epsilon}), \ \epsilon > 0$$

In chapter 3, I show that if $h_n = n^{-\alpha}$ for some $\alpha > 0$, the conditions for the relative rate of convergence is the following:

$$\begin{cases} 
\alpha < \frac{1}{3}, \\
\alpha \leq \frac{1}{2} \delta, \\
\alpha > \frac{1}{2M}.
\end{cases}$$

Thus,

$$\hat{R}(t) - R(t) = \frac{1}{n} \sum_{i=1}^n \psi_i^{R,1}(\psi_i^\beta, t) + \frac{1}{nh_n} \sum_{i=1}^n \psi_i^{R,2}(Y_i, X_i, t; t_n) + \xi_n^R(t),$$

where by construction $E[\psi_i^{R,1}(\psi_i^\beta, t)] = E[\psi_i^{R,2}(Y_i, X_i, t; t_n)] = 0$ and $\sup_{t \in \mathcal{T}} |\xi_n^R(t)| = O_p \left( n^{-\frac{1}{2} - \epsilon} \right)$, for some $\epsilon > 0$.

Combined with assumptions (3.3.1), (3.3.2) and (3.3.3), it completes the proof of theorem (3.3.1).

C.2 Reminding terms in $\hat{S}$

Two reminding terms in (3.8):

$$\xi_n^{S,2} \equiv \int R(t) \left( 1\{\hat{R}(t) \geq -b_n\} - 1\{R(t) \geq 0\} \right) dQ(t) +$$

$$+ \int \left( \hat{R}(t) - R(t) \right) \left( 1\{\hat{R}(t) \geq -b_n\} - 1\{R(t) \geq 0\} \right) dQ(t)$$

(i) First, I show that $\sup_{t \in \mathcal{T}} 1\{|\hat{R}(t) - R(t)| > b_n\} = o_p(n^{-\epsilon})$ for any $\epsilon > 0$.

Assume the following two classes of functions are Euclidean with a constant envelope (more pritive conditions?):

$$\mathcal{R}_1 = \{\psi^{R,1}(\cdot, t) : t \in \mathcal{T}\}, \quad \mathcal{R}_2 = \{\psi^{R,2}(\cdot, \cdot, t; h_n) : t \in \mathcal{T}, h_n > 0\}.$$
Define the following two empirical processes $\nu_{R_1,n}(\cdot)$ and $\nu_{R_2,n}(\cdot)$ over $\mathcal{F}$:

$$
\nu_{R_1,n}(t) = \frac{1}{n} \sum_{i=1}^{n} \psi^{R_1}(\psi_i^\beta, t), \quad \nu_{R_2,n}(t) = \frac{1}{n} \sum_{i=1}^{n} \psi^{R_2}(Y_i, X_i, t; h_n).
$$

Thus, there exists a constant $\bar{D} > 0$ such that for any $\delta > 0$,

$$
\mathbb{P} \left[ \sup_{t \in \mathcal{F}} \left| \sqrt{n} \nu_{R_1,n}(t) \right| \geq \delta \right] \leq \frac{\bar{D}}{\delta}, \quad \mathbb{P} \left[ \sup_{t \in \mathcal{F}} \left| \sqrt{n} \nu_{R_2,n}(t) \right| \geq \delta \right] \leq \frac{\bar{D}}{\delta}.
$$

By the uniform properties of $\xi_n^R(t)$, there exists $\bar{C}$ such that for any $\delta > 0$:

$$
\mathbb{P} \left[ \sup_{t \in \mathcal{F}} \left| \xi_n^R(t) \right| > \delta \right] \leq \frac{\bar{C}}{\delta n^{\frac{1}{2}+\varepsilon}}, \quad \text{for some } \varepsilon > 0.
$$

Since,

$$
\hat{R}(t) - R(t) = \nu_{R_1,n}(t) + \frac{1}{h_n} \nu_{R_2,n}(t) + \xi_n^R(t),
$$

$$
\mathbb{P} \left[ \sup_{t \in \mathcal{F}} \left| \hat{R}(t) - R(t) \right| > b_n \right] \leq \mathbb{P} \left[ \sup_{t \in \mathcal{F}} \left| \nu_{R_1,n}(t) \right| > \frac{b_n}{3} \right] + \mathbb{P} \left[ \sup_{t \in \mathcal{F}} \left| \frac{1}{h_n} \nu_{R_2,n}(t) \right| > \frac{b_n}{3} \right] + \mathbb{P} \left[ \sup_{t \in \mathcal{F}} \left| \xi_n^R(t) \right| > \frac{b_n}{3} \right] < \frac{3\bar{D}}{b_n n^{\frac{1}{2}}} + \frac{3\bar{D}}{b_n n^{\frac{1}{2}} h_n} + \frac{3\bar{C}}{b_n n^{\frac{1}{2}+\varepsilon}}.
$$

Requirements for the rate of convergence:

$$
\begin{align*}
&\begin{cases}
  b_n n^{\frac{1}{2}} \to \infty, \\
  b_n n^{\frac{1}{2}} h_n \to \infty, \\
  b_n n^{\frac{1}{2}} + \varepsilon \to \infty.
\end{cases}
\quad \iff \begin{cases}
  b_n n^{\frac{1}{2}} \to \infty, \\
  b_n n^{\frac{1}{2}} h_n \to \infty.
\end{cases}
\end{align*}
$$

Consider a random variable $1 \left\{ \sup_{t \in \mathcal{F}} \left| \hat{R}(t) - R(t) \right| \geq b_n \right\}$. Show that it is $o_p(n^{-\varepsilon})$. For large enough $n$ and any $\varepsilon > 0$ and $\delta > 0$,

$$
\mathbb{P} \left[ n^\varepsilon \left\{ \sup_{t \in \mathcal{F}} \left| \hat{R}(t) - R(t) \right| \geq b_n \right\} > \delta \right] = \mathbb{P} \left[ \sup_{t \in \mathcal{F}} \left| \hat{R}(t) - R(t) \right| \geq b_n \right]
$$
The last term converges to zero with $n \to \infty$. So, $1\left\{\sup_{t \in \mathcal{I}} |\hat{R}(t) - R(t)| \geq b_n\right\} = o_p(n^{-\epsilon})$ for any $\epsilon > 0$.

(ii) Second, assume that there exists some $D > 0$ such that

$$\lim_{b \downarrow 0} \frac{P_Q\{-2b \leq R(t) < 0\}}{b} \leq D.$$ 

Thus, $\int 1\{-2b_n < R(t) < 0\}dQ(t) = O_p(b_n)$.

(iii) To analyze $1\{\hat{R}(t) \geq -b_n\} - 1\{R(t) \geq 0\}$, consider four cases:

(a) $\hat{R}(t) \geq -b_n$ and $R(t) \geq 0$: thus, both indicators equal 1 and the difference between them is 0

(b) $\hat{R}(t) \geq -b_n$ and $R(t) < 0$: thus, the first indicator is 1 and the second is 0, the difference is 1. Note that

$$P[\hat{R}(t) \geq -b_n \text{ and } R(t) < 0] \leq P[|\hat{R}(t) - R(t)| \geq b_n]$$

(c) $\hat{R}(t) < -b_n$ and $R(t) \geq 0$: thus, the first indicator is 0 and the second is 1, the difference is -1. Note that,

$$P[\hat{R}(t) > -b_n \text{ and } R(t) \leq 0] \leq P[-2b_n < R(t) < 0] + P[|\hat{R}(t) - R(t)| \geq b_n]$$

(d) $\hat{R}(t) < -b_n$ and $R(t) < 0$: thus, both indicator equal 0 and the difference is 0

Use results from (i)-(iii) and assuming that $\int |R(t)|dQ(t)$ is bounded:

$$\left|\int R(t) \left(1\{\hat{R}(t) \geq -b_n\} - 1\{R(t) \geq 0\}\right) dQ(t)\right| \leq$$

$$\leq \int |R(t)| \left|1\{\hat{R}(t) \geq -b_n\} - 1\{R(t) \geq 0\}\right| dQ(t) \leq$$

$$\leq \int |R(t)|(1\{-2b_n < R(t) < 0\} + 2 \cdot 1\{|\hat{R}(t) - R(t)| \geq b_n\})dQ(t) \leq$$

$$\leq O(b_n^2) + o_p(n^{-\epsilon}).$$

If $b_n^2n^{-\frac{1}{2}} \to 0$, then the first term in $\xi_n^{S,2}$ is of order $O_p(n^{-\frac{1}{2}-\epsilon})$, for some $\epsilon > 0$.  

126
(iv) From (C.4), by definition:

$$\sup_{t \in T} \left| \hat{R}(t) - R(t) \right| = O_p \left( \frac{1}{n^{\frac{1}{2}} h_n} \right).$$

The second term in $\xi_n^{S,2}$ is:

$$\left| \int \left( \hat{R}(t) - R(t) \right) \left( 1\{\hat{R}(t) \geq -b_n\} - 1\{R(t) \geq 0\} \right) dQ(t) \right| = O_p(n^{-\frac{1}{2}} h_n^{-1})o_p(n^{-\epsilon}) = o_p(n^{-\frac{1}{2}-\epsilon} h_n^{-1})$$

Thus,

$$|\xi_n^{S,2}| = O_p(n^{-\frac{1}{2}-\epsilon}).$$


Vita

Lidia Kosenkova

Education

• Ph.D. in Economics, Pennsylvania State University, 2018
  Fields: Econometric Theory, Applied Econometrics, Industrial Organization
  Committee: Andrés Aradillas-López, Joris Pinkse, Marc Henry, Bing Li

• M.A. in Economics, New Economic School, 2012

• B.S. in Mathematics and Computer Science, Moscow State University, 2009

Employment

• Assistant Professor of Economics, University of Virginia, 08/2018 — present

Publications

Working Papers

• “Nonparametric Inference in Asymmetric First-price Auctions with $k$-rationalizable Beliefs” (Job Market Paper) (PDF)

• “Inference in First-price Auctions with Endogenous Entry under $k$-rationalizable Beliefs”

Conferences and Seminars

• 2018: University of Calgary, Baruch College, Instituto Tecnológico Autónomo de México, Rensselaer Polytechnic Institute, New York University, University of Virginia, University of Technology Sydney, Australian National University, University of Sydney, University of California San Diego, University of Illinois at Urbana-Champaign, Southern Economic Association 88th Annual Meeting (scheduled)

• 2017: University of Western Ontario, Pennsylvania State University, Cornell University

Last updated: June 25, 2018