TCER Working Paper Series

A Distribution-Free Test of Monotonicity with an Application to Auctions

Yusuke Matsuki

September 2016

Working Paper E-110
http://tcer.or.jp/wp/pdf/e110.pdf

TOKYO CENTER FOR ECONOMIC RESEARCH
1-7-10-703 Iidabashi, Chiyoda-ku, Tokyo 102-0072, Japan

Abstract

Abstract This study develops a simple distribution-free test of monotonicity of conditional expectations. The test is based solely on ordinary least squares (OLS) and exploits the property between conditional expectation and projection; we prove that the monotonicity of a conditional expectation function restricts the sign of a corresponding projection coefficient. The estimated projection coefficient is used for a one-tailed t-test. The test -- which is notably simpler than other monotonicity tests -- is applied to bidding data from Japanese construction procurement auctions to empirically test first-price sealed bid auction models with independent private values (IPV), assuming the data are generated from a symmetric Bayesian Nash equilibrium. We regress the bid level on the number of bidders and use the estimated projection coefficient for testing. We find that the test results depend on public work categories.

Yusuke Matsuki
Nagasaki University
Faculty of Economics
4-2-1 Katafuchi, Nagasaki 850-8506, Japan
ymatsuki@nagasaki-u.ac.jp

A Distribution-Free Test of Monotonicity with an Application to Auctions

Yusuke Matsuki*

September 3, 2016

Abstract

This study develops a simple distribution-free test of monotonicity of conditional expectations. The test is based solely on ordinary least squares (OLS) and exploits the property between conditional expectation and projection; we prove that the monotonicity of a conditional expectation function restricts the sign of a corresponding projection coefficient. The estimated projection coefficient is used for a one-tailed t-test. The test — which is notably simpler than other monotonicity tests — is applied to bidding data from Japanese construction procurement auctions to empirically test first-price sealed bid auction models with independent private values (IPV), assuming the data are generated from a symmetric Bayesian Nash equilibrium. We regress the bid level on the number of bidders and use the estimated projection coefficient for testing. We find that the test results depend on public work categories.

1 Introduction

In economic theory, it is fundamental to distinguish the economic model describing reality from both positive and normative perspectives. For instance in auction theory, a bidder's strategy depends on the auction's environment and its rules. Thus, the predictions of auction theory – including the optimal design of auctions – depend on the modeling framework. For instance, if an auction environment is specified with the common value (CV) paradigm, policies may be introduced by the auction designer to minimize the impact of the "winner's curse." Even when the private values assumption

*Faculty of Economics, Nagasaki University, 4-2-1 Katafuchi, Nagasaki 850-8506, Japan (email: ymatsuki@nagasaki-u.ac.jp). Financial supports from Tokyo Center for Economic Research, Nomura Foundation, and JSPS KAKENHI (Grant Number 16K21231) are gratefully acknowledged.

is maintained, bidders may be less aggressive as the number of bidders increases, assuming bidder valuations are affiliated in the sense of Milgrom and Weber (1982). Deciding upon the applicable framework leads to improved policy and strategy advice.

Economists have derived testable implications of models to investigate the extent to which the integrity of a proposed model is maintained when presented with actual data. Monotonicity is a prominent qualitative feature of certain economic models such as auction models.

In this paper, we develop a simple distribution-free test of monotonicity of conditional expectations. The test is based solely on ordinary least squares (OLS) and exploits the property between conditional expectation and projection which states that the monotonicity of the conditional expectation function restricts the sign of a projection coefficient. The test statistic is the t-value of an estimated projection coefficient and we perform a one-tailed t-test, which is notably simpler than other monotonicity tests.

Only a few studies or textbooks discuss the relational properties of these two concepts in spite of their fundamental role in statistical analysis.[1] The relation between conditional expectation functions and projection coefficients is not simple because of the functional form of the projection coefficients. However, if the monotonicity of a conditional expectation function is established, the sign of a corresponding projection coefficient represents the slope of the conditional expectation function.

Furthermore, we apply our test to data from Japanese procurement auctions to investigate whether auction models of independent private values (IPV) can justify the data. In procurement auctions, the bidder who submits the lowest bid wins the contract. We consider a benchmark model of auction theory in which bidders are risk neutral and engage in a symmetrical Bayesian Nash equilibrium (BNE) strategy, which we call the standard IPV model. The model has a monotone relation between bids and the number of bidders. We also consider an extension in which the set of bidders is uncertain when the actual bidders form their bid. We refer to the standard IPV model with this extension as the IPV model.

We assume that data are generated from the IPV model, and we empirically investigate the relation between bids and the number of bidders by regressing the bid level on the number of bidders in an auction. We focus on selective tendering, in which bidders are selected by the auctioneer, to avoid considering strategic bidder entry. In general, there is no fixed relationship between the bid level and the number of bidders. However, in the standard IPV model, the average bid must decrease

[1] Wooldridge (2002), Angrist and Pischke (2009), and Hansen (2016) provide a lengthy argument on this topic. Our discussion in Section 2.2 below parallels the one in these books.

as the number of bidders increases.[2] After controlling for auction heterogeneity, a positive coefficient for the number of bidders indicates the rejection of the standard IPV model. A negative coefficient indicates that the standard model survives and might yet potentially justify the observations.

The test uses a reduced form equation related to an equilibrium bid function. The analysis is nonparametric, i.e., there is no parametric assumption on the distribution of bidder costs and the form of conditional expectations regarding bids. Reduced form coefficients have no exact correspondence to the structural model but are in some sense sufficient to represent the IPV models.

Tests of monotonicity have been studied in the last twenty years. Ellison and Ellison (2011) exploits the monotone relationship to investigate strategic entry deterrence. Gutknecht (2016) employs a second-order U-process to develop a consistent test. For other works in the literature, see the reference in Chetverikov (2013) and Gutknecht (2016). Our test is much simpler and allows multiple explaining and omitted variables at the cost of a limitation described later.

Previous surveys by Hendricks and Paarsch (1995) and Laffont (1997) classify the empirical analyses of auctions into theory testing and structural estimations of auction models.[3,4] The literature on theory testing has a longer history but has been thin in recent years, whereas the structural estimation literature has expanded as the result of methodological developments and increases in computational resources (see Laffont et al. (1995), Guerre et al. (2000)).

As Athey and Haile (2007) note, although various theoretical implications that can be used to construct formal tests have been derived, there are only scant test studies. The research by Hendricks et al. (2003) is a careful and persuasive positive study questioning whether game theory can explain real bidding behavior, but their tests are not formal in the sense of statistical inference. On the other hand, we can perform a *formal* test.

The number of bidders has been considered in the context of common value models in which focus has been placed on the “winner’s curse.” In common value auctions, winning suggests that the winner overestimated the true value of the object. Moreover, if the winner is myopic and does not consider this update, the winner loses benefits, a phenomenon that is referred to as the “winner’s curse.” However, auction theory assumes that rational bidders take the winner’s curse into account and bid

[2] Note that this is a statement about buying by the auctioneer such as procurement. In the case of selling, a larger set of bidders leads to higher bids.

[3] Another approach comparable to the auction theory approach is the decision theoretic approach.

[4] Note that in principle, both are complementary. We can develop various specification tests based on statistics obtained by structural estimation.

smaller amounts, thus implying that the bid level tends to fall as the number of bidders increases, i.e., auctions become less competitive. This prediction is sharp, counter-intuitive and seemingly easy to test with available data and a simple regression analysis.

To the best of our knowledge, this study is the first to test IPV models using the number of bidders and every bid from the auctions. The literature testing IPV models is extremely limited compared with studies regarding the winner's curse. Empirical auction research should be interested in the extent to which the benchmark IPV framework is maintained because identification and estimation procedures are well established in IPV models when proceeding to structural analysis. Moreover, this study is the first to use a one-sided t-test for auctions.

Related auction studies include Gilley and Karels (1981), Brannman et al. (1987), Thiel (1988), Levin and Smith (1991) and Hendricks et al. (2003), while the more recent literature includes Boone and Mulherin (2008) and Tse et al. (2011). Gilley and Karels (1981) applied a decision theory approach to develop a statistical model and controlled the participation of potential bidders by employing a sample selection correction technique. These authors tested the "pure" common values model with the data from an outer continental shelf (OCS) petroleum lease sale, revealing that bidders are less competitive as their numbers increase. Brannman et al. (1987) appears to be the only study investigating the relation between the number of bidders and the bid level in the context of the IPV model. These authors used only winning bids, whereas we include all bids. The use of only winning bids is too restrictive for testing the IPV model. Moreover, we provide a formal interpretation by establishing the relation between the projection coefficients and the conditional expectation function.

In OCS auctions, it is natural to specify auctions with common value models and employ the Hendricks and Porter framework (Hendricks et al. (1994, 2003)). However, in procurement auctions, it is less clear which model should be used (e.g., Hong and Shum (2002), Thiel (1988)). A handful of studies employ the IPV framework, although recent works extensively consider entry.

Guerre et al. (2000) and Athey and Haile (2002, 2007) propose more tests of the IPV model that rely on structural estimates, such as the estimated distribution of bidder valuations. However, our simple test serves as an initial tool to validate the standard IPV auction model.

The remainder of this study is organized as follows. In Section 2, we discuss the monotone relationship between conditional expectation and projection. In Section 3, our test is developed. Section 4 reviews auction theory related to this research. Section 5 explains the data. Section 6 presents the econometric model and the test.

The empirical results are presented in Section 7. Section 8 concludes.

2 Monotonicity of Conditional Expectation and Projection

In this section, we discuss the relation between conditional expectation and projection. We first focus on the univariate case for simplicity but the argument below is easily extendable to the multivariate case (see Section 3.1.). Let y and x be random variables satisfying the assumption of the finite second moment: $E[x^2] < \infty$ and $E[y^2] < \infty$ (and invertible covariance matrix in the multivariate case). The projection of y on x, denoted as $L[y|x]$, is written as follows:

$$L[y|x] \equiv \gamma_0 + \gamma_1 x$$

where γ_0 and γ_1 are the projection coefficients. There exists a unique projection error η such that

$$y = L[y|x] + \eta$$

and $E[\eta] = E[x\eta] = 0$. Let $E[y|x]$ be the conditional expectation of y on x. There exists an unique error ε such that

$$y = E[y|x] + \varepsilon$$

and $E[\varepsilon] = E[\varepsilon|x] = 0$. The only one that relates the two is the following:

$$L[y|x] = L[E[y|x]|x]. \tag{1}$$

We reinterpret this law of iteration from the point of view of the linear approximation of functions.

Let $m(x) \equiv E[y|x]$. We consider a strictly increasing m as a benchmark case but the argument is easily applicable to nondecreasing/increasing or strictly decreasing cases. We prove that if m is strictly increasing, then $\gamma_1 > 0$. This result is intuitively clear but not easy to prove in the case of multiple explanatory variables as is shown below. If we assume the linearity in conditional expectation, i.e., $E[y|x] = \alpha_0 + \alpha_1 x$, the statement is clear, but this imposes a severe restriction on the joint distribution of (x, y). Essentially, no assumption on the distribution is imposed if only projection is considered.

2.1 Linear Approximation of a Function

Definition 1. Let f be a function to be approximated. Let $g(x) = \gamma_0 + \gamma_1 x$ be a linear function. Then, g is said to be a linear approximation of f over $[a, b]$ with respect to F if (γ_0, γ_1) minimize

$$\int_a^b (f(x) - g(x))^2 dF(x) = \int_a^b (f(x) - \gamma_0 - \gamma_1 x)^2 dF(x)$$

where F is a nondecreasing and right-continuous function. Moreover, g is said to be a (global) linear approximation of f if g is a linear approximation of f over the domain of f.

This indicates that the distance between a function f and a linear function g is minimized. This is a special case of least squares approximation of functions. We can contrast this with a Taylor approximation, which is essentially a local approach with tangent lines. In the linear approximation, we take an average over $[a, b]$.

2.2 Conditional Expectation and Projection

Next, we review the result that the projection $L[y|x]$ is a linear approximation of the conditional expectation $E[y|x]$. Recall that the coefficients $\gamma = (\gamma_0, \gamma_1)$ in $L[y|x] = \gamma_0 + \gamma_1 x_1$ are a solution to

$$\min_{(c_0, c_1)} E[(y - c_0 - c_1 x)^2].$$

On the other hand, from equation (1),

$$L[y|x] = L[E[y|x]|x] = L[m(x)|x].$$

Thus, γ is a solution to

$$\min_{(c_0, c_1)} E[(m(x) - c_0 - c_1 x)^2] = \min_{(c_0, c_1)} \int (m(x) - c_0 - c_1 x)^2 dF(x)$$

where F is the cumulative distribution function of x. Hence, the projection $L[y|x]$ is a linear approximation of the conditional expectation $E[y|x]$ with respect to F.

2.3 Monotonicity

We prove the key original result stating that if $m(x)$ is strictly increasing in x, then $\gamma_1 > 0$. An algebraic approach does not work to prove this property because $\gamma =$

$E[\mathbf{x}\mathbf{x}']^{-1}E[\mathbf{x}y]$ ($\gamma = (\gamma_0, \gamma_1, \dots \gamma_k)'$ and $\mathbf{x} = (1, x_1, \dots x_k)'$) is complicated in the multivariate case ($k \geq 2$) and it is not clear how to use the formula (a system of linear equations) to prove $\gamma_1 > 0$. If there is only one explanatory variable ($k = 1$), the result may be easy. The formula in this case is

$$\gamma = \begin{pmatrix} \gamma_0 \\ \gamma_1 \end{pmatrix} = \begin{pmatrix} E[y] - \frac{Cov(x,y)}{Var(x)} E[x] \\ \frac{Cov(x,y)}{Var(x)} \end{pmatrix}.$$

If you establish the covariance $Cov(x, y)$ is positive, then $\gamma_1 > 0$. However, if $k \geq 2$, we cannot rely on this approach. Let $\mathbf{x} = (1, x_1, x_2)'$ and suppose we need to prove $\gamma_1 > 0$. Then

$$\gamma = \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \end{pmatrix} = E \begin{bmatrix} 1 & x_1 & x_2 \\ x_1 & x_1^2 & x_1 x_2 \\ x_2 & x_1 x_2 & x_2^2 \end{bmatrix}^{-1} \begin{pmatrix} E[y] \\ E[x_1 y] \\ E[x_2 y] \end{pmatrix}.$$

As is observed from the formula above, γ_1 has a complicated form and it is difficult to utilize this expression for the proof. We rely on a more visual and intuitive approach using basic calculus.

Theorem 1. *Let f be a strictly increasing function defined on $[a, b]$. Let $g(x) = \gamma_0 + \gamma_1 x$ be a linear function. If g is a linear approximation of f, $\gamma_1 > 0$.*

Proof. The first order conditions with respect to γ_0 and γ_1 yield the two equations below:

$$\int_a^b (f(x) - \gamma_0 - \gamma_1 x) dF(x) = 0$$

and

$$\int_a^b x(f(x) - \gamma_0 - \gamma_1 x) dF(x) = 0.$$

If $\gamma_1 \leq 0$, there exists $c \in [a, b]$ such that $f(x) - g(x) \leq 0$ for $[a, c]$ and $f(x) - g(x) \geq 0$ for $[c, b]$. From the first equation,

$$\int_a^c (f(x) - g(x)) dF(x) + \int_c^b (f(x) - g(x)) dF(x) = 0.$$

Then, given $f(x) - g(x) > 0$ for $x > c$,

$$\begin{aligned}
\int_a^b x(f(x) - g(x))dF(x) &= \int_a^c x(f(x) - g(x))dF(x) + \int_c^b x(f(x) - g(x))dF(x) \\
&= k_1 \int_a^c (f(x) - g(x))dF(x) + k_2 \int_c^b (f(x) - g(x))dF(x) \\
&= k_1 \left(- \int_c^b (f(x) - g(x))dF(x) \right) + k_2 \int_c^b (f(x) - g(x))dF(x) \\
&= (k_2 - k_1) \int_c^b (f(x) - g(x))dF(x) \\
&> 0
\end{aligned}$$

where k_1 and k_2 are constants satisfying $a < k_1 < c < k_2 < b$ whose existence is guaranteed by the first mean value theorem for integration. This inequality contradicts the second equation of the first order conditions. □

The same argument can be applied to the nonincreasing or nondecreasing cases or the strictly decreasing case. Finally, we have the following result.

Theorem 2. *If the conditional expectation $E[y|x]$ is nondecreasing (nonincreasing) in x, the projection coefficient of x is nonnegative (nonpositive). If the conditional expectation $E[y|x]$ is strictly increasing (decreasing) in x, the projection coefficient of x is positive (negative).*

3 Hypothesis Testing of Monotonicity

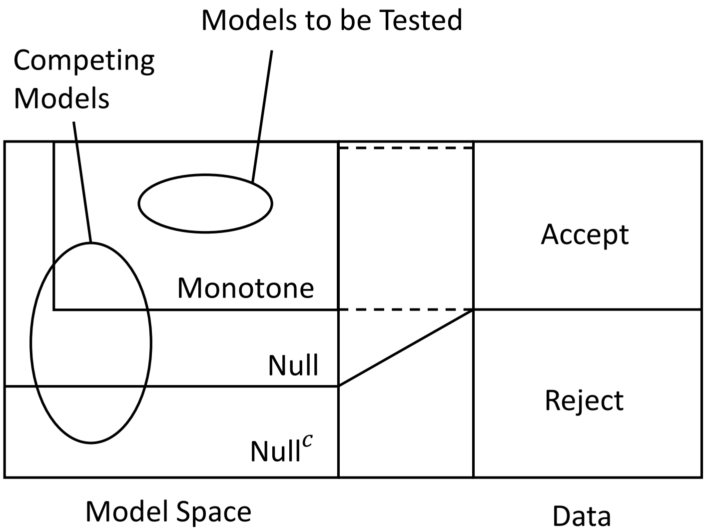
The result above is used to construct a test of monotonicity. Again, let $m(x) \equiv E[y|x]$ and $\gamma_0 + \gamma_1 x = L[y|x]$. If m is strictly increasing, then $\gamma_1 > 0$, and the test is specified below:

$$H_0 : \gamma_1 > 0 \quad \text{versus} \quad H_1 : \gamma_1 \leq 0.$$

The null is that the coefficient of the variable of interest is positive.[5] Rejection of the null implies rejection of the hypothesis that the function is strictly increasing. "Acceptance" of the null implies the possibility of the conditional expectation function being strictly increasing.

[5]If $m(x)$ is nondecreasing, then γ_1 is nonnegative, and the test is $H_0 : \gamma_1 \geq 0 \quad H_1 : \gamma_1 < 0$. If $m(x)$ is strictly decreasing, then γ_1 is negative, and the test is $H_0 : \gamma_1 < 0 \quad H_1 : \gamma_1 \geq 0$. If $m(x)$ is nonincreasing, then γ_1 is nonpositive, and the test is $H_0 : \gamma_1 \leq 0 \quad H_1 : \gamma_1 > 0$.

Figure 1: Conceptual diagram of models and tests



This is only a one-tailed t-test. The information required to implement this test is the estimate of projection coefficient γ_1 with its standard error. Let $\widehat{\gamma_1}$ be the projection estimate of γ_1, $SE[\widehat{\gamma_1}]$ be the standard error of $\widehat{\gamma_1}$, and $t[\widehat{\gamma_1}] = \widehat{\gamma_1}/SE[\widehat{\gamma_1}]$ be the t-value. The p-value of the test is given by $p[\widehat{\gamma_1}] = \Phi\left(t[\widehat{\gamma_1}]\right)$, where Φ is the CDF of the standard normal distribution. For example, with a significance level $\alpha = 0.05$, if $t[\widehat{\gamma_1}] < -1.645$, the null is rejected.

In short, once we establish the monotonicity of conditional expectations, we can test the monotonicity with a one-tailed t-test. This is *the* simplest formal test for monotonicity.

We discuss the implication of the test result with its limitation. The implication is obvious in rejection. The function cannot be strictly increasing. On the contrary, "acceptance" adds nothing in general, which is the difference between our test and other consistent (asymptotic power 1) tests. The function *can* be strictly increasing, but the results provide no additional information about distinguishing the two. For our purpose of testing auction models, however, this limitation is not a crucial fault, as we discuss later.

The argument above is more easily understandable with a Venn diagram. See Figure 1. The left box is the set of models which can possibly explain the data. "Models to be Tested" is the models of interest to be tested. "Monotone" is the set of models which implies monotonicity. "Null" is the set of models which implies the null hypothesis. "Null^c" is the complement of "Null". The right box named "Data" is the range of test statistics which has "Acceptance" and "Rejection" region. For example, if the null of a test is $H_0 : \gamma_1 > 0$ with a significance level $\alpha = 0.05$, "Data"$= \mathbb{R}$, "Rejection"$= (-\infty, -1.645)$, and "Acceptance"$= [-1.645, \infty)$.

Suppose the set of competing models lies in "Monotone", "Null", and "Null^c". In our test, the correspondence between the models and the regions of the test statistic is expressed with the solid lines. If the value of the test statistic is in the "Acceptance" region, the true model is in "Null". While the models of interest can be true, the true model needs not be monotone and can be a competing model. If the data yields rejection, the true model is in "Null^c" and the model of interest cannot be true. With a consistent test, the correspondence between the models and the regions of the test statistic is expressed with the dashed lines. If the value of the test statistic is in the "Acceptance" region, the true model must be monotone. Note that competing models which also imply monotonicity can be true. In the case of rejection, the true model is in the complement of "Monotone" and the model of interest cannot be true.

3.1 Multivariate Case

Our test is extendable to a multivariate case paralleling the above argument. Let $x = (x_1, \ldots, x_A)$ and $z = (z_1, \ldots, z_B)$. Suppose $f(x, z)$ is strictly increasing in each x_j for all x_{-j} and z, where $x_{-j} = (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_A)$. Then, a linear approximation of f is presented as

$$g(x, z) = \gamma_0 + \gamma_1 x_1 + \gamma_2 x_2 + \cdots + \gamma_A x_A + \gamma_{A+1} z_1 + \cdots + \gamma_{A+B} z_B$$

and $\gamma_j > 0$ $(1 \leq j \leq A)$. With this expression, $E[y|x, z]$ and $L[y|x, z]$ have the same relation as the two in the univariate case, and we have a test

$$H_0 : \forall j \in \{1, \ldots, A\}\ \gamma_j > 0 \quad \text{and} \quad H_1 : \exists j \in \{1, \ldots, A\}\ \gamma_j \leq 0.$$

The test statistic can be constructed from $\{\gamma_j\}_{(1 \leq j \leq A)}$ (See Kodde and Palm (1986)).

4 Auction Models and Implications

We apply our test to specification testing of the IPV auction model. Before proceeding to the application, we review auction models to derive a testable implication.

Let $\{1, 2, \ldots, n\}$ be the set of bidders where $n \geq 2$. An indivisible contract is auctioned in a first-price sealed bid auction in which the lowest bidder wins the contract. Let $x \in [0, \overline{x}]$ be the privately known cost of completing the contract for a bidder. The private cost of each bidder is drawn from a common distribution whose absolutely continuous cumulative distribution function is denoted as F. In other words, bidders are ex-ante symmetric. In addition, the structure of the game is common knowledge among bidders.

We focus on a standard symmetric Bayesian Nash equilibrium. The equilibrium bid β is given by

$$\beta(x;n) = x + \int_x^{\overline{x}} \frac{(1-F(t))^{n-1}}{(1-F(x))^{n-1}} dt.$$

Taking the difference with respect to the number of bidders n, we have

$$\beta(x;n+1) - \beta(x;n) = \int_x^{\overline{x}} \left(\frac{1-F(t)}{1-F(x)} - 1 \right) \frac{(1-F(t))^{n-1}}{(1-F(x))^{n-1}} dt.$$

Because

$$\frac{1-F(t)}{1-F(x)} - 1 \leq 0$$

for $x \leq \overline{x}$, and $x \leq t$, we conclude that

$$\beta(x;n+1) < \beta(x;n)$$

for all $x \in [0, \overline{x}]$ and $n \geq 2$. This finding supports the intuition that more competitors lead to more aggressive bids. Hence, $F_{B|n}$, the cumulative distribution function of bids conditional on the number of bidders being n, first-order stochastically dominates $F_{B|n+1}$ and

$$E[B|n] > E[B|n+1]$$

for all $n \geq 2$, where B is a bidder's bid. The test proposed in this study relies on the latter property. If bid B is regressed on the number of bidders n, the coefficient should be negative. This is a testable implication of the standard first-price auction model with IPV.

4.1 Unknown Number of Bidders

In addition to the standard model, we also consider the unknown number of bidders model because in the application, we use data that contain (possibly) both types of auctions. Some auctions are conducted with the set of bidders hidden prior to bidding. However, it is impossible to distinguish based on the data which auctions disclosed the list of bidders prior to bidding.

The study of auctions with the unknown number of bidders was started by McAfee and McMillan (1987) and Harstad et al. (1990) and generalized by Myerson (1998) and Milchtaich (2004). We use the simple property that when the number of bidders is not known to each other, then bidder's strategy cannot depend on the number of bidders. We follow Harstad et al. (1990) for the description of the unknown number of bidders model.

The number of actual bidders n_a is uncertain and stochastic at the time of bidding in this case. Let $p(n_a)$ be the ex ante probability of the number of actual bidders being n_a. Each bidder is assumed to be selected symmetrically; in other words, each potential bidder is selected with an equal probability conditional on n_a. The probability that the number of other actual bidders is $\widetilde{n_a}$ for an actual bidder $\widetilde{p}(\widetilde{n_a})$ is

$$\widetilde{p}(\widetilde{n_a}) = \frac{n_a p(n_a)}{\sum_{k=1}^{N} kp(k)} \quad (n_a = \widetilde{n_a} + 1)$$

where N is the number of potential bidders. The unique symmetric equilibrium strategy is as follows:

$$\beta^*(x) = \sum_{k=1}^{n} w_k(x)\beta(x; k)$$

where β is given before and

$$w_k(x) = \frac{F(x)^{k-1}\widetilde{p}(k)}{\sum_i F(x)^{i-1}\widetilde{p}(i)}.$$

Note that β^* does not depend on the number of actual bidders.

4.2 Unobserved Heterogeneity

Unobserved heterogeneity has received much attention in recent years (e.g., Arcidiacono and Miller (2011), Shiu and Hu (2013)). In empirical auction analysis, Krasnokutskaya (2011, 2012) consider a specific functional form for the observable and unobservable components of auctions. Roberts and Schlenker (2013), Aradillas-López et al. (2013), Armstrong (2013), and Hu et al. (2013) consider a more general functional form. Unobserved heterogeneity in auction is omitted variables which are observed by the bidders but not by the econometrician. We discuss an omitted variable problem in a general context.

Let x be the variable of interest, z be the other observable conditioning variables including the intercept and w be the omitted variable.[6] Suppose $E[y|x, z, w]$ is strictly increasing in x for all z and w. The law of total expectations yields that

$$E[y|x = x_0, z = z_0] = \int E[y|x = x_0, z = z_0, w = w_0] f_{w|x,z}(w_0|x_0, z_0) dw_0$$

[6] We assume here that x and w are scalar random variables and z is a vector of random variables, but we are able to consider a more general case of all variables being a vector.

where $f_{w|x,z}$ is the PDF of w conditional on (x, z). Though an increase in x shifts the integrand $E[y|x, z, w]$ upward, this also changes the weight $f_{w|x,z}$ and the monotonicity in x does not hold in general. Hence, our test cannot be applied without additional assumptions or data. One possible condition to guarantee the monotonicity is the coincidence of the shift direction of $E[y|x, z, w]$ and $f_{w|x,z}$: if $E[y|x, z, w]$ is increasing in w for all (x, z) and $E[y|x, z, w]$ and $f_{w|x,z}$ move in the w-direction as x increases for all z, this would imply the monotonicity of $E[y|x, z]$ in x.

Another approach is a more basic one using the omitted variables argument in OLS estimation. Let the projection of y on (x, w, z) be

$$y = \gamma_x x + z'\gamma_z + \gamma_w w + \varepsilon. \qquad (\gamma_x > 0)$$

On the other hand, the projection of y on (x, z) is written as follows:

$$y = \alpha_x x + z'\alpha_z + \eta.$$

Then

$$\begin{pmatrix} \alpha_x \\ \alpha_w \end{pmatrix} = \begin{pmatrix} \gamma_x \\ \gamma_w \end{pmatrix} + E\left[\begin{pmatrix} x \\ w \end{pmatrix}\begin{pmatrix} x \\ w \end{pmatrix}'\right]^{-1} E\left[\begin{pmatrix} x \\ w \end{pmatrix} z'\right]\gamma_z.$$

If x is uncorrelated with z, $\gamma_x = \alpha_x$ and our test can be applied. If x is correlated with z, we can use instrumental variables to correct for the omitted variables bias.

4.3 Other Models and Implications of the Test

When we perform descriptive (reduced form) analysis, we should be clearer about what can be said based on the sign of coefficient estimates. It is generally difficult to formally test specifications only from reduced form analysis in auction theory. One exception is basic IPV models.

Previous studies that have empirically investigated common value auction models typically use the number of bidders and the bid level to test the winner's curse effect. However, this approach tests the models in an informal way. Note that even in common value models, the magnitude of the discouraging winner's curse effect cannot be easily determined. Although theory suggests that bids tend to be less aggressive as the number of bidders increases, the competition effect may prevail, and increasing the number of bidders in some cases leads to more aggressive bids.

Even in the private value models, bidders can be less aggressive as the number of bidders increases when values are affiliated, as shown by Pinkse and Tan (2005). Thus, rejection in the test does not necessarily imply that affiliated private values

(APV) models are inappropriate.

5 Data

We employ publicly available data from the Kanto Regional Development Bureau (KRDB) of the Ministry of Land, Infrastructure, Transport and Tourism (MLIT). The Japanese Regional Development Bureaus are responsible for managing transportation and water infra-structure, parks, and public buildings in the nation.[7] They hold auctions for contracts of various types of construction, maintenance, and repair projects involving roads, bridges, and public buildings. The KRDB covers Ibaraki, Tochigi, Gunma, Chiba, Tokyo, and Kanagawa prefectures, and portions of Nagano and Shizuoka prefectures.

The data on public works contract procurement auctions were obtained from the KRDB. The data for each contract auction include the contract name, bid date, category, the name office which offers the contract, and engineer's estimate of the contract, in addition to the bidder names and bids. We use selective tendering data from 2005, in which the Japanese central government employed selective tendering for procurement auctions of construction contracts for smaller public works contracts.[8] The system subsequently moved toward an open bid format, and selective tendering drastically decreased after 2005.

We restrict our attention to selective tendering because we must account for the strategic entry of bidders in open tendering. In open tendering, we must distinguish between potential bidders and actual bidders, and potential bidders typically cannot be observed or are not easy to identify. In models with strategic entry, the relation between the number of actual bidders observed and the bid level may not be monotone.[9] By focusing on selective tendering, entry is much less important and the possible model is closer to the benchmark, which enables us to test the IPV model more accurately.

Table 1 reports summary statistics for the number of bidders and engineer's estimates of auctions. Table 2 is a frequency table of the number of bidders. Although

[7]Japan has eight Regional Development Bureaus, including the KRDB. For a more detailed description of their work, see http://www.kkr.mlit.go.jp/en/.

[8]We focus on "pure" selective tendering ("tsujo shimei kyoso nyusatsu" in Japanese). In this format, bidders who are not selected never bid. In addition, there are other selection formats involving two-step selection processes.

[9]Note that even when the set of potential bidders is known to the econometrician, the test cannot be applied because the relation between the number of potential bidders and the bid level may not be monotone. Li et al. (2009) demonstrate that if a model considers strategic entry, bidders may become less aggressive as the number of potential bidders increases, even in the IPV framework.

these numbers range from 5 to 24, approximately 95% of auctions have 10 to 13 bidders. The MLIT formally defines 21 categories for public works procurement while the data include 18 construction categories with auctions and bids. Summary statistics for the total number of contracts and bids for each category are included in Table 3. We focus on the categories in which the total number of bids exceeds 300.

6 Econometric Model and Testing

Now we describe our econometric model for testing. We include auction-specific covariates z_t of auction t described below, and we assume that bidder valuations are drawn from a conditional cumulative distribution function $F(\cdot|z_t)$. The econometric model is

$$B_{tk} = \gamma_0 + \gamma_{num} numBids_t + \gamma_{est} est_t + \gamma_{off} office_t + \gamma_{mon} month_t + \eta_{tk} \tag{2}$$

where B_{tk} is the bid of bidder k, $numBids_t$ is the number of bidders, est_t is the engineer estimate, $office_t$ is the managing office, $month_t$ is the month, and η_{tk} is the projection error for auction t. The explaining variables constitute z_t. Note that z_t has no bidder subscript because we consider the symmetric auction model and bidder specific information is not used. Because we consider projection, $E[z_t \eta_{tk}] = 0$ but $E[\eta_{tk}|z_t] \neq 0$ in general. $E[B_{tk}|z_t]$ is arbitrary and not generally linear. Equation (2) is a linear approximation of the expected equilibrium bid conditional on z_t.

The standard IPV model implies that $\gamma_{num} < 0$. If the set of bidders is uncertain, a bidder's strategy cannot depend on the number of bidders. If we control auction-specific covariates, the average bid does not depend on the realization of the number of bidders, which means $\gamma_{num} = 0$. Overall, the test can be organized as follows:

$$H_0 : \gamma_{num} \leq 0 \quad H_1 : \gamma_{num} > 0.$$

If H_0 is not rejected, the IPV model, which allows both cases of a known and unknown set of bidders, is not rejected in and could justify the data. If H_0 is rejected, the IPV model cannot be justified based on the data and must be excluded from our consideration.

We discuss implications from the test result as for the empirical analysis of auctions. We compare two tests, one of which is ours and the other of which is a consistent test such as the one by Gutknecht (2016). If the null is rejected in both tests, this implies that the IPV model is not true with a same significance level. However, if the null is not rejected, this implies in our test that the IPV model survives as a candidate to explain the data, but other competing models like the CV and APV models also remain candidates. The "acceptance" result implies in the other test that the hypothesis of decreasing function cannot be supported by the data because the test has a power of 1, and we can more convincingly believe in the nondecreasing relationship. However, the argument stops there. Note that in the CV or APV models, the nondecreasing relationship is possible. In both tests, "acceptance" means the IPV model is possible, but other possibilities like the CV or APV models cannot be excluded. Thus our test is sufficient for our purpose. In Figure 1, the model of interest is the IPV model and the competing models include the CV and APV models.

7 Estimation Results

Using the above equation, we regress the bid level on the number of bidders, the cost estimates, and dummies for offices and months in each selected public works category. The OLS estimation results for each category are included in Table 4. Our interest is in the columns "NumBids" and "p-value." The "NumBids" column reports the estimate of coefficient γ_{num}, whereas "p-value" reports the p-value of the test.

We find that the estimation results vary by category. The coefficient of the engineer estimate is significantly positive – as expected – but the values differ. The coefficients of the number of bidders vary considerably, and even the signs are reversed. In over half of the selected categories, the coefficient for the number of bidders is not significant. This result may imply that the set of bidders is not disclosed in most auctions in these categories. In categories 1 (public engineering), 5 (construction engineering), 12 (coating work) and 20 (communication equipment installation), the coefficients are significant. In categories 1, 5 and 12, the coefficients are significantly negative, which implies that the IPV model is maintained. The bidders become aggressive as their numbers increase. Our main focus is the p-value of the test. In categories 4 (landscape gardening), 19 (machinery installation) and 20, the hypothesis is rejected at the 10% significance level. In addition, the null is strongly rejected, particularly in category 20.

The results imply that in these categories, the IPV model cannot be maintained. The rejection implies that affiliated signals or common value components may explain the results, or a violation of competitive bidding, such as colluding behavior, may have occurred. Another possible explanation when assuming common value components and the unknown set of bidders is that auction covariates are positively correlated with the number of bidders and that bidders are less aggressive because of the high number of bidders expected. Moreover, there might have been unobserved auction heterogeneity which reversed the sign of the true projection coefficient of the number of bidders.

Finally, there is one caveat that must be stated regarding this interpretation of the results. In Japanese procurement, the engineer's estimate is treated as a secret reserve price, whereby bids above it never win the contract and the value is hidden prior to bidding. If we wish to evaluate the result in a more rigorous manner, we need to establish the monotone relationship in the model of the secret reserve price. Nonetheless, it is reasonable to conjecture that the monotone relationship holds in this case and the empirical analysis provides meaningful results.

8 Conclusions

We have established a structural interpretation of projection coefficients and developed a simple distribution free test of monotonicity. We have applied our test to data from Japanese procurement auctions to investigate whether auction models of IPV can justify the data and find that in some public works categories, the IPV model is rejected. The proposed test is simple and based solely on OLS, and it does not require a specific restriction regarding the cost distribution.

The limitation of the approach in this study is that the test does not deeply exploit the structure of the model. Thus, if the null is not rejected, there are many other possible models and it is necessary to continue exploring to distinguish the model's structure. In our application, CV and APV models are possible regardless of the test results. Unobserved heterogeneity may complicate the interpretation of the results. Nonetheless, due to its simplicity, our test could serve as an initial tool for an empirical analysis with both the structural estimation and the test of the theory.

References

Angrist, J. D. and J.-S. Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton: Princeton University Press.

Aradillas-López, A., A. Gandhi, and D. Quint (2013). Identification and inference in ascending auctions with correlated private values. *Econometrica 81*(2), 489–534.

Arcidiacono, P. and R. A. Miller (2011). Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica 79*(6), 1823–1867.

Armstrong, T. B. (2013). Bounds in auctions with unobserved heterogeneity. *Quantitative Economics 4*(3), 377–415.

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

Athey, S. and P. A. Haile (2007). Nonparametric approaches to auctions. In J. J. Heckman and E. E. Leamer (Eds.), *Handbook of Econometrics*, Volume 6, Part A, Chapter 60. Amsterdam: Elsevier.

Boone, A. L. and J. H. Mulherin (2008). Do auctions induce a winner's curse? new evidence from the corporate takeover market. *Journal of Financial Economics 89*(1), 1–19.

Brannman, L., J. D. Klein, and L. W. Weiss (1987). The price effects of increased competition in auction markets. *Review of Economics and Statistics 69*(1), 24–32.

Chetverikov, D. (2013). Testing regression monotonicity in econometric models. Technical report, UCLA.

Ellison, G. and S. F. Ellison (2011, September). Strategic entry deterrence and the behavior of pharmaceutical incumbents prior to patent expiration. *American Economic Journal: Microeconomics 3*(1), 1–36.

Gilley, O. W. and G. V. Karels (1981). The competitive effect in bonus bidding: New evidence. *The Bell Journal of Economics 12*(2), 637–648.

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

Gutknecht, D. (2016). Testing for monotonicity under endogeneity: An application to the reservation wage function. *Journal of Econometrics 190*(1), 100–114.

Hansen, B. E. (2016). *Econometrics.* available online at http://www.ssc.wisc.edu/~bhansen/econometrics/Econometrics.pdf.

Harstad, R. M., J. H. Kagel, and D. Levin (1990). Equilibrium bid functions for auctions with an uncertain number of bidders. *Economics Letters 33*(1), 35–40.

Hendricks, K. and H. J. Paarsch (1995). A survey of recent empirical work concerning auctions. *The Canadian Journal of Economics / Revue canadienne d'Economique 28*(2), 403–426.

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

Hendricks, K., R. H. Porter, and C. A. Wilson (1994). Auctions for oil and gas leases with an informed bidder and a random reservation price. *Econometrica 62*(6), 1415–1444.

Hong, H. and M. Shum (2002). Increasing competition and the winner's curse: Evidence from procurement. *Review of Economic Studies 69*(4), 871–898.

Hu, Y., D. McAdams, and M. Shum (2013). Identification of first-price auctions with non-separable unobserved heterogeneity. *Journal of Econometrics 174*(2), 186–193.

Kodde, D. A. and F. C. Palm (1986). Wald criteria for jointly testing equality and inequality restrictions. *Econometrica 54*(5), 1243–1248.

Krasnokutskaya, E. (2011). Identification and estimation of auction models with unobserved heterogeneity. *Review of Economic Studies 78*(1), 293–327.

Krasnokutskaya, E. (2012). Identification and estimation of auction model with two-dimensional unobserved heterogeneity. *International Economic Review 53*(3), 659–692.

Laffont, J.-J. (1997, January). Game theory and empirical economics: The case of auction data. *European Economic Review 41*(1), 1–35.

Laffont, J.-J., H. Ossard, and Q. Vuong (1995, July). Econometrics of first-price auctions. *Econometrica 63*(4), 953–80.

Levin, D. and J. L. Smith (1991). Some evidence on the winner's curse: Comment. *American Economic Review 81*(1), 370–375.

Li, S., C. Timmins, and R. H. von Haefen (2009, September). How do gasoline prices affect fleet fuel economy? *American Economic Journal: Economic Policy 1*(2), 113–37.

McAfee, R. P. and J. McMillan (1987). Auctions with a stochastic number of bidders. *Journal of Economic Theory 43*(1), 1–19.

Milchtaich, I. (2004). Random-player games. *Games and Economic Behavior 47*(2), 353–388.

Milgrom, P. R. and R. J. Weber (1982). A theory of auctions and competitive bidding. *Econometrica 50*(5), 1089–1122.

Myerson, R. B. (1998). Population uncertainty and poisson games. *International Journal of Game Theory 27*, 375–392.

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

Roberts, M. J. and W. Schlenker (2013). Identifying supply and demand elasticities of agricultural commodities: Implications for the us ethanol mandate. *American Economic Review 103*(6), 2265–95.

Shiu, J.-L. and Y. Hu (2013). Identification and estimation of nonlinear dynamic panel data models with unobserved covariates. *Journal of Econometrics 175*(2), 116–131.

Thiel, S. E. (1988). Some evidence on the winner's curse. *American Economic Review 78*(5), 884–895.

Tse, M., F. Pretorius, and K. Chau (2011). Market sentiments, winner's curse and bidding outcome in land auctions. *The Journal of Real Estate Finance and Economics 42*, 247–274.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge: MIT Press.

Table 1: Summary statistics

	min	mean	s.d.	median	max	N
Number of bidders	5	11.51	2.16	11	24	1032
Engineer estimates	2.86	39.70	25.76	32.86	96.00	1032

Notes: "Engineer estimates" is measured in millions of Japanese yen.

Table 2: Frequency of the number of bidders

Number of bidders	Freq.	Percent
≤ 9	15	1.45
10	372	36.05
11	237	22.97
12	145	14.05
13	223	21.61
$14 \leq$	40	3.88
Total	1,032	100

Notes: This table reports the frequency of the number of bidders in auctions. The number of bidders that are equal to or smaller than 9, and equal to or larger than 14 are aggregated into "≤ 9" and "$14 \leq$", respectively.

Table 3: Total number of contracts and bids

Cat. No.	Cat. Name	Num. Cont.	Num. Bids
1	Public engineering	177	2011
2	Paving work	34	410
3	Steel bridge superstructure work	3	35
4	Landscape gardening	56	638
5	Construction engineering	103	1225
6	Wooden building construction	0	0
7	Electrical work	36	426
8	Piping work	27	332
9	Cement concrete work	1	13
10	Prestressed concrete work	2	21
11	Slope surface treatment	8	98
12	Coating work	39	576
13	Maintenance and repair work	396	4466
14	Dredging work	5	53
15	Grouting work	0	0
16	Piling work	0	0
17	Well drilling work	1	13
18	Prefabricated building construction	2	24
19	Machinery installation	41	421
20	Communication equipment installation	80	888
21	Receiving and transforming facility installation	21	233

Notes: This table presents the total number of contracts and bids of selective tenders for each category in 2005. “Cat. No.” indicates category number, which we assigned, and “Num. Cont.” indicates the number of contracts.

Table 4: OLS estimation results

Category	NumBids	S.E.	Engineer est.	S.E.	p-value	N
1	-0.320*	0.178	0.867***	0.0221	0.964	2,011
2	-0.241	0.615	0.730***	0.107	0.653	410
4	0.294	0.200	0.920***	0.0272	0.0706	638
5	-0.993***	0.294	0.849***	0.0271	1	1,225
7	0.341	0.620	1.062***	0.0602	0.291	426
8	1.487	4.302	0.907***	0.179	0.365	332
12	-0.662*	0.360	0.777***	0.0453	0.967	576
13	0.0291	0.103	0.846***	0.0113	0.389	4,466
19	1.644	1.146	0.428***	0.0860	0.0757	421
20	2.565***	0.610	0.444***	0.0524	1.29E-05	888

Notes: ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Standard errors are reported in the "S.E." column. "NumBids" stands for the coefficient of the number of bidders for each auction, and "Engineer est." stands for the coefficient of the reserve price. "p-value" indicates the p-value for the one-tailed t-test of $H_0 : \gamma_{num} \leq 0$ against $H_1 : \gamma_{num} > 0$. For all categories, months and offices are controlled with dummy variables.