Série des Documents de Travail

n° 2019-15

Distinguishing Incentive from Selection Effects in Auction-Determined Contracts

Laurent LAMY[1]
Manasa PATNAM[2]
Michael VISSER[3]

[1] CIRED, Ecole des Ponts, ParisTech. E-mail : laurent.lamy78@gmail.com
[2] IMF. E-mail : mpatnam1@gmail.com
[3] CREST; ENSAE; CRED, University of Paris 2. E-mail : michael.visser@ensae.fr

Distinguishing Incentive from Selection Effects in Auction-Determined Contracts *

Laurent Lamy[†] Manasa Patnam [‡] Michael Visser[§]

September 20, 2019

Abstract

This paper develops a novel approach to estimate how contract and principal-agent characteristics influence an ex-post performance outcome when the matching between agents and principals derives from an auction process. We propose a control-function approach to account for the endogeneity of contracts and matching. This consists of, first, estimating the primitives of an interdependent values auction model - which is shown to be non-parametrically identified from the bidding data - second, constructing control functions based on the distribution of the unobserved private signals conditional on the auction outcome. A Monte Carlo study shows that our augmented outcome equation corrects well of the endogeneity biases, even in small samples. We apply our methodology to a labor market application: we estimate the effect of sports players' auction-determined wages on their individual performance.

Keywords: Econometrics of Contracts, Econometrics of Auctions; Structural Econometrics; Endogenous Matching; Polychotomous Sample Selection; Wage-Performance Elasticity.
JEL classification: C01; C29 ; C57; D44; M52; Z22.

*We thank James Albrecht, Shekar Shankar Aiyar, Mitali Das, Kartikeya Date, Xavier d'Haultfoeuille, Aureo de Paula, Guido Kuersteiner, Ingmar Prucha, Paul Scott, Ali Skalli, Andrew Sweeting, Frank Vella and seminar participants at CREST, Georgetown University, the University of Besançon, the University of Luxembourg, the University of Maryland, the University of Paris II, the 2nd ECODEC Conference, the 7th French Econometrics Conference, and the 46-th EARIE conference, for helpful comments and suggestions. We thank Arun Kumar for excellent research assistance. This paper is a major revision of a draft which circulated under the title "Correcting for Sample Selection From Competitive Bidding, with an Application to Estimating the Effect of Wages on Performance." Financial support from the LABEX ECODEC (ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047) is gratefully acknowledged.

†CIRED (Ecole des Ponts, ParisTech). Email: laurent.lamy78@gmail.com.
‡IMF. Email: mpatnam1@gmail.com.
§ENSAE-CREST and CRED, University of Paris 2. Email: michael.visser@ensae.fr.

1 Introduction

A central issue in the empirical literature on contracts concerns quantifying the effect of contract characteristics on observed behavior. These characteristics influence ex-post outcomes through two kinds of channels: on the one hand, they offer a variety of incentives to elicit efforts from the contracting parties; on the other, they play a key role in the way agents and principals match together according to both observable and unobservable characteristics. For instance, to reduce the cost of monitoring, a firm can decide to raise salaries, which, by making the unemployment threat more costly, should incite workers to exert more effort (Shapiro and Stiglitz (1984)). However, the wage increase also changes the pool of workers that are matched to the firm, in particular on the basis of unobserved worker characteristics, including those that create idiosyncratic synergies between workers and firms (Abowd et al. (1999)). Similarly, recent empirical work on procurement contracts awarded through competitive tendering have stressed the importance of both ex-post moral hazard and adverse selection (Lewis and Bajari (2011), Decarolis (2014)): agents' ex-post performance (like the project-delivery time or the quality of realized work) is driven by the contract characteristics (part of which may be endogenously determined at the procurement auction)[1] both directly through incentives schemes (like penalties for delays) and indirectly through the fact that contract characteristics have a screening effect (e.g., the ability of a firm to complete the project on time may be correlated with its winning probability). An assessment of the relative magnitude of incentive and adverse selection effects is crucial when deriving either optimal contracting choices or optimal market regulation (Laffont and Tirole (1993)).

In the empirical contract literature, most papers focus on the "incentive effect" of contracts. These effects are typically identified using exogenous variations in contract characteristics, by way of field or natural experiments. Such an approach is particularly popular in the labor literature, e.g., to estimate whether better paid workers reciprocate (Gneezy and List (2006), Mas (2006), Lee and Rupp (2007)), or to measure to what extent salaries on the basis of realized output (like piece-rates wages or sharecropping arrangements) enhance performance ceteris paribus (Lazear (2000), Shearer (2004), Burchardi et al. (2019)).

Although quantifying incentive effects is admittedly interesting, it is by itself not sufficient to answer many important policy questions. If one wishes to measure for instance the impact on a principal's payoff of using an alternative class of contracts, or to evaluate the welfare effect of a new regulation, then counterfactual calculations are necessary. Such calculations require in turn that the distribution of agents' unobserved characteristics is known. They furthermore require knowledge of how self-selection of agents to contracts is determined given these unobserved variables. Distinguishing incentive from selection effects is, however, a challenging

[1] In the so-called 'A+B' design analyzed by Lewis and Bajari (2011), construction firms not only submit a dollar bid (for materials and labor) but also a duration in days to complete the project which determines how penalties for delays are computed.

issue as stressed by Chiappori and Salanié (2003).

This issue has been addressed when a monopolistic principal posts a contract (or a menu of contracts) to the agents who decide first whether to match with the principal and second chose their level of effort or consumption, and this optimally given their privately observed characteristics. To identify such models, two kinds of approaches have been developed in this literature: One strand relies on exogenous variations of the contracts posted by the principal (Powell and Goldman (forthcoming), d'Haultfoeuille and Février (forthcoming)). Another assumes that the contract posted by the principal is optimal and relies then explicitly on the corresponding first-order conditions. The seminal theoretical contribution in this field is Perrigne and Vuong (2011) who establish the non-parametric identification of Laffont and Tirole (1986)'s procurement model from the observation of the contract's realization (the quantity demanded, the production costs and the associated transfers). Perrigne and Vuong (2011) show how to distinguish in the realized production costs what comes from the ex-ante type of the agent (capturing the effect of adverse selection), the agent's effort (driven by the incentives associated to the chosen contract), and the ex-post exogenous shocks affecting final demand and (observed) production costs.[2] While this literature considers typically uni-dimensional types as in the bulk of auction and contract theory, some recent contributions have dealt with contract models involving multidimensional types (Luo et al. (2012), Aryal et al. (2019)).

In many environments, however, there are multiple principals to be matched with the various agents. This paper develops a new methodology to consistently estimate how contract and principal-agent characteristics influence agents' ex-post performance when the matching of agents to principals and part of the contract characteristics are determined through auctions.[3] We present our methodology through the labor auction framework that we use later as an empirical illustration: to buy the services of workers, firms participate to a centralized auction market that assigns each worker to the winning bidder (if any) and where the winning bid determines the wage.[4] Nevertheless, we should keep in mind that our methodology can be applied very broadly.

The main features of our setup are the following: An ex-post measure of each worker's performance is observed. Besides firm-worker observable characteristics, this measure is assumed to depend on the wage (the incentive effect) and also on two signals. The first signal reflects idiosyncratic worker-firm synergies, observed privately by the firm concerned, and the second one reflects attributes that are commonly valued by all firms, observed by the firm previously

[2]See Paarsch and Shearer (2000) and Gayle and Miller (2015) for the estimation of related agency models applied to labor contract data, and Perrigne and Vuong (2011) for additional references, in particular regarding contractual arrangements in agriculture, corporate finance, credit markets, insurance and retailing.

[3]We consider thus a class of models where the matching process involves "market power". This contrasts with the frictionless two-sided matching models where there is a large set of players on both sides, and which are popular to model marriage markets with unobserved heterogeneity (Choo and Siow (2006)). The econometrics of such matching models has been surveyed by Chiappori and Salanié (2016).

[4]Labor markets wherein employers and job-seekers are matched through an auction proliferate on the Internet. See `bit.ly/OnlineLaborAuction` for examples.

employing the worker, if there is one, and referred to as the incumbent. Each firm is assumed to value the worker through the future performance and in addition through two signals that are unrelated to performance, one idiosyncratic signal and one commonly valued signal observed by the incumbent. This full set of signals drives then the auction outcome, i.e., determines whether the worker is hired by a firm, and if so the wage and the identity of the firm the worker is matched with. Our model structure implies that 1) a match between an agent and a principal reflects idiosyncratic synergies between the two; 2) the final wage which results from competitive bidding is driven not only by the winning bidder's signals but also by the losing bidders' signals; 3) the commonly valued signals drive the incumbent's bid and hence the wage, and this even if the incumbent quits before the auction terminates. The commonly valued signals are a source of asymmetry between the incumbent and the other bidders in terms of bidding behavior and hence matching probabilities. Beyond auctions, features 1) and 2) are expected to show up in most models with imperfection competition,[5] while feature 3) is a novel aspect of our model.

Our setting involves thus endogenous selection of contract features based on multidimensional unobserved private signals and an inter-linkage of selection and incentive effects. To tackle the resulting endogeneity biases in the performance equation, we propose a novel methodology which involves two steps. First, we analyze the bidding stage where workers are auctioned through a sequence of English auctions and where firms are assumed to anticipate correctly workers' future performance during the bidding process. Concretely, we assume that when deciding to remain active or not in the auction, firms perfectly anticipate that higher wages may induce higher (or lower) performance. They also update their beliefs according to the incumbent's bid. Our auction model involves therefore interdependent values with asymmetric bidders and furthermore departs from the "quasi-linear payoff" paradigm that typically prevails in the auction literature. After characterizing the perfect Bayesian equilibrium, the main result of this first step consists in showing that our auction model -and in particular the underlying signal distributions that drive bidding behavior- are non-parametrically identified from the observation of the winning price and the winner's identity alone. Our identification result does not require the existence of exogenous variables characterizing the worker and/or the firms, nor the presence of binding reserve prices, but relies rather on three main restrictions: first, the independence of the signals across bidders; second, the independence between the privately and the commonly valued signals; and third, an exclusion restriction stipulating that the distributions of the unobserved attributes are identical in sub-samples with and without incumbents. While the first independence assumption is needed to identify the pure private values model (Athey and Haile (2007)), the last two restrictions are crucial to deal with an interdependent value model.[6]

[5] It arises e.g. in Laffont and Tirole (1987)'s optimal procurement model and in Berry et al. (1995)'s differentiated products market model.

[6] Analogs of our second independence restriction are used to deal with unobserved heterogeneity thanks to a deconvolution argument (Li et al. (2000) and Krasnokutskaya (2011)). Restrictions that are similar to our exclusion

Second, given identification of our auction model, Bayesian updating conditional on the observables allows us to identify the distribution of the bid-signals appearing in the performance equation up to a finite set of parameters. This in turn allows us to build and identify two types of control functions: One type is associated with the privately valued idiosyncratic synergy signals of the winning bidder, while the other is associated with the commonly valued signals received by the incumbent. Once inserted into the post-auction outcome equation, these control terms account for the various sources of endogeneity, and we can then consistently estimate the incentive of the wage and the impact of firm-worker observable characteristics (including the incumbency indicator) on the post-auction outcome. A notable novel aspect of our control terms is their dependence not only on the characteristics of the matched worker-firm pair but also on the characteristics of the losing bidders, which is valuable to reduce potential problems of collinearity between our control terms and the explanatory variables.

A Monte Carlo study shows that estimation of the post-auction outcome equation by OLS (i.e., ignoring endogeneity) leads to strongly biased estimates of the coefficients on both the principal-agent characteristics and the contract incentive, and this even in the absence of any incentive effect (i.e., when wages have no effect on the post-outcome variable) or when there is no incumbent (i.e., when the auction model is a pure private value one). Furthermore, the extent of these biases tends to grow with the strength of the commonly valued signal present in the auction. Our endogeneity-corrected estimator is, however, shown to be unbiased even in relatively small samples.

We illustrate our methodology using data regarding the English auctions determining cricket players' assignment and corresponding wages for a major tournament held each year in India, paired with post-auction data regarding (sold) players' performance. More than 300 cricket players are "auctionned" through a centralized broadcasted bidding market and about a third of them are "bought" and then assigned to one of the 8 teams participating in the tournament. A first objective is to determine how wages affect performance. The effect is found to be statistically significant and positive, both using OLS and our method. OLS leads, however, to a substantially larger estimated effect than our method, which means that naive estimation of the performance equation results in an upward bias in the incentive effect. A second objective is to estimate whether players perform better (or not) when staying in the same team, i.e., whether being hired by the incumbent represents an advantage (or a disadvantage) in terms of performance. The effect is found to be negative but insignificant using OLS while our method leads to a much larger and statistically significant negative effect. This discrepancy is consistent with the fact that endogeneity plays an important role in our data, especially because our control terms are highly correlated with the incumbency indicator.

Our work contributes to the recent structural empirical literature that considers the joint modeling of auction and post-auction data. In the seminal contributions of Athey and Levin (2001)

restriction are used in the literature to test for common values (Haile et al. (2003)), to identify risk-aversion (Lu and Perrigne (2008)), or to account for correlated private values (Aradillas-Lopez et al. (2013)).

and Hendricks et al. (2003), the post-auction data consists of the characteristics of the item auctioned (proportion of different timber species and quantity of oil, respectively), on which bidders have only imperfect information at the auction stage. Unlike our case their post-auction outcome is therefore exogenous: in particular it is influenced neither by the characteristics of the auction winner nor by the bids.[7] More recently, Bhattacharya et al. (2018) and Bodoh-Creed and Hickman (2018) have dealt, similarly to us, with endogenous post-matching data where the matching process between agents and principals is modeled with an auction model. Although their auction models differ from ours, they share with us the fact that the post-auction and auction stages are structurally linked. Bhattacharya et al. (2018) analyze jointly bidding and drilling data from a set of oil companies, through the lens of a symmetric pure common value model where each company receives for each tract a unidimensional private signal about the quantity of oil. The post-auction outcome (drilling decision) only depends on contract features and the oil quantity which is assumed to be revealed after the auction, in contrast with our model wherein the post-auction outcome depends on the identity of the winner and where bidders are asymmetrically informed about commonly valued attributes. Bodoh-Creed and Hickman (2018) model college enrollment as an all-pay auction where students (who are eager to enter the best colleges) compete in terms of an academic score which represents a sunk cost that depends on the student (unobserved) learning ability and the efforts he/she made. This matching model allows the authors to analyze a post-graduation performance equation and to disentangle the impact of college quality from students' learning abilities and efforts, and also to evaluate the impact of affirmative action through various counterfactual exercises.[8]

We also contribute to the literature that uses a control function approach to correct for selectivity and/or endogeneous explanatory variables (see Vella (1998) and Wooldridge (2015) for surveys). Heckman (1979)'s seminal contribution, extended to polychotomous outcomes by Dubin and McFadden (1984), deals with selectivity only. A novelty of our paper is that we deal with these two issues simultaneously and with a non-parametric perspective.[9] Furthermore, unlike the vast majority of articles in this field that adopt a reduced-form approach, our control functions are constructed by explicitly relying on (auction) theory.

The remainder of the paper is organized as follows. Section 2 presents the environment and discusses how the omission of bidders' private signals can bias OLS estimates. Section 3 characterizes how firms and workers are matched as the equilibrium of an interdependent auction model. Section 4 shows that this model is non-parametrically identified from the bidding data

[7] As argued by Athey and Haile (2007), such post-auction information allows to depart from the pure private value paradigm that prevails due to identification limitations.

[8] Relatedly, Cuesta and Sepúlveda (2018) analyze the repayment rate equation for consumer credits. They model the matching between banks and borrowers as the outcome of a process where each borrower decides whether to incur or not a sunk cost in order to organize a second-price auction in which banks compete in terms of credit rates.

[9] Wooldridge (2002) and Das et al. (2003) also propose a control function method to address jointly these two issues with a parametric and non parametric approach, respectively. The setup considered by those authors is not easily generalized to allow for selection on multiple outcomes, and in particular to take into account the way sample selection and endogeneity arise in our auction setup.

alone, which enables us in turn to identify the control functions. Section 5 relates our contribution to the literature and discusses identification under alternative assumptions. Section 6 presents the results of some Monte Carlo simulations. Section 7 describes the labor market environment used for our application and empirical results. Section 8 concludes and discusses other possible applications of our methodology.

2 The environment and the endogeneity problem

We consider a (small) collection of firms, indexed by $f = 1, \dots, F$, bidding for a (large) collection of workers, indexed by $i = 1, \dots, N$. The workers are auctioned sequentially, one after another. An observable reserve price, denoted by $W_i^r \geq 0$, is attached to each worker i before the auction sequence starts: it corresponds to the starting price in the auction for i. The winner of this auction (if there is one) is denoted f_i^w, and the wage at which i is employed corresponds to the (observable) auction price, and is denoted $w_i \geq W_i^r$. The bidding rules are detailed in the next section and the associated equilibrium will then determine both the auction price and the auction winner as a function of bidders' private information. We take into account the possibility that i worked in one of the firms prior to the auctions. In such a case, the corresponding firm is referred to as the incumbent and denoted by $f_i^{inc} \in \{1, \dots, F\}$. Otherwise we let $f_i^{inc} := 0$.

Before the auction for worker i starts, each firm f receives a single-dimensional private signal, denoted by $s_{i,f}^{PV} \in \mathbb{R}$. This signal is only observed by f and summarizes match-specific attributes regarding the worker's future performance if hired by f (e.g., synergies between the managerial staff and the worker). The incumbent associated to i (if any) receives also an additional single-dimensional private signal, denoted by $s_i^{CV} \in \mathbb{R}$. This signal is only observed by the incumbent f_i^{inc} and summarizes attributes relative to i's future performance that are commonly valued by all firms (e.g., the worker's ability). It captures determinants which were revealed to the incumbent through earlier interaction with this worker. If i is employed by firm f, then we consider that it is the sum of the private signals $s_{i,f}^{PV}$ and s_i^{CV} that matters in determining i's future performance. The incumbent thus has superior information as in Engelbrecht-Wiggans et al.'s (1983) pure common value auction model with an informed bidder.[10] Formally, the performance of worker i if employed by firm f, denoted $y_{i,f} \in \mathbb{R}$, is assumed to take the following form:

$$y_{i,f} = \beta_f + \beta_x \cdot x_{i,f} + \beta^{inc} \cdot x_{i,f}^{inc} + \tau \cdot \log(w_i) + \underbrace{s_{i,f}^{PV} + s_i^{CV} + \epsilon_{i,f}}_{:=u_{i,f}} \tag{1}$$

[10]Beyond the auction and procurement literature, the incumbency status has received some attention in the labor literature: in particular, insider information accruing to incumbents is shown to impede employer-to-employer worker flows (Greenwald (1986)) and significantly affect the distribution of equilibrium wages as a function of seniority (Pinkston (2009)). See Oyer and Schaefer (2011)'s survey for more details on the impact of the presence of incumbents in the labor market.

where the term $\tau \cdot \log(w_i)$ captures the direct effect of wage on performance and corresponds thus to the incentive effect referred to in the introduction.[11] The parameter β_f represents a firm-specific fixed effect, $x_{i,f}^{inc}$ the dummy variable that is equal to one if $f = f_i^{inc}$ and zero otherwise, β^{inc} the parameter measuring the effect of the incumbency indicator on performance (e.g., capturing idiosyncratic gains from match-specific experience), $x_{i,f}$ a vector of observable characteristics of both i and f that are publicly observable before the auction starts, β_x a vector of parameters measuring the effects of the corresponding characteristics.[12] Finally, $\epsilon_{i,f}$ captures other performance determinants that are unobserved by the firms. We assume that the vectors $(\epsilon_{i,1}, ..., \epsilon_{i,F})$ are i.i.d. across all i, that they are drawn independently of all other variables in our model, and that their means are equal to zero. The signals s_i^{CV} and $s_{i,f}^{PV}$ are also centered around zero (additional assumptions on these signals are introduced in Section 3). The econometrician neither observes these two signals nor the performance shock. From his/her perspective the error term in the performance equation is the sum of these three components, denoted $u_{i,f}$.

In our analysis, the presence of incumbents is a source of two kinds of asymmetries between firms: 1) an informational asymmetry at the bidding stage due to the signal s_i^{CV} which is only observed by the incumbent; 2) a performance asymmetry due to the term $\beta^{inc} \cdot x_{i,f}^{inc}$ appearing in the performance equation. Disentangling the two is challenging from an econometric perspective: that is the reason why almost all structural empirical works on auctions with an incumbent are limited to a single of these sources of asymmetry.[13]

To estimate the performance equation, we face two kinds of problems. The first is that estimation is based on a selected sample. The selection arises because the wage and performance of a given worker i are only observed if some firms decide to participate, i.e., to enter a bid at or above the reserve price W_i^r. In addition, if there is bidding at or above the reserve price, only recorded is the performance y_{i,f_i^w}, while $y_{i,f}$ is naturally unknown and counterfactual for all $f \neq f_i^w$. Given that we expect the incumbent firm f_i^{inc} (resp. a non-incumbent firm $f \neq f_i^{inc}$) to bid more aggressively the higher the term $\beta_{f_i^{inc}} + \beta_x \cdot x_{i,f_i^{inc}} + \beta^{inc} + s_{i,f_i^{inc}}^{PV} + s_i^{CV}$ (resp. $\beta_f + \beta_x \cdot x_{i,f} + s_{i,f}^{PV}$) – as shown formally in the next section – the sample selection is not random but related to $x_{i,f}$, $x_{i,f}^{inc}$ and $s_{i,f}^{PV}$ (for all f) and to s_i^{CV}. Belonging to the selected sample or not thus depends on the observed firm-worker characteristics, but also on the unobserved signals received by the firms. Both these types of variables are at the same time determinants of performance. The estimation of the performance equation (1) by OLS leads to biased estimates because of the link between the selection rule and the composed

[11] Note that the incentive effect does not vary across workers. If the effect were worker-dependent and anticipated by firms/bidders, then we would be in a setup where moral hazard drives selection which is typically excluded, a notable exception being Einav et al. (2013).

[12] Note that $x_{i,f}$ may include contract characteristics established prior to the auction in which case the associated elements in β_x represent incentive effects as well.

[13] See Hendricks and Porter (1988), Hendricks et al. (1994) and Pinkston (2009) for pure common value models with informational asymmetry, and De Silva et al. (2009) and Allen et al. (2019) for pure private value models with cost/valuation asymmetries between the incumbent and the entrants. By combining the two forms of asymmetries, Weiergraeber and Wolf (2018) is a notable exception.

error term $u_{i,f}$. More precisely, there is a bias in the OLS estimates because the mean of this error term conditional on being in the sample (and given the wage and the other observable explanatory variables), which is denoted by $E[u_{i,f_i^w}|w_i, x_{i,f_i^w}, x_{i,f_i^w}^{inc}]$, is non-zero and varies across observations.[14] For example, if the incumbent wins the auction (i.e., $f_i^w = f_i^{inc}$), we have $E[u_{i,f_i^w}|w_i, x_{i,f_i^w}, x_{i,f_i^w}^{inc}] = E[s_{i,f_i^w}^{PV} + s_i^{CV}|w_i, x_{i,f_i^w}, x_{i,f_i^w}^{inc}]$, which is expected to decrease with $\beta_{f_i^w} + \beta_x \cdot x_{i,f_i^w} + \beta^{inc}$ (if this term is small the winner only bids aggressively if the sum of the signals $s_{i,f_i^w}^{PV} + s_i^{CV}$ is large). If β^{inc} is positive (resp. negative), the OLS estimate of this parameter is then expected to be downward (resp. upward) biased.[15] As a consequence, a standard OLS analysis may lead the analyst to wrongly conclude that there are no synergies between worker i and the incumbent f_i^{inc}, while the true value of β^{inc} differs from zero. Note that sample selection causes the estimator of β_x to be biased as well, including the parameters associated with variables $x_{i,f}$ that do not depend on f (such as contract features set prior to the auction). Again, we expect the OLS estimate of β_x to be biased toward zero.

The second problem is related to the first one, but is nonetheless distinct. It concerns the fact that w_i is potentially endogenous in (1). Large values of $u_{i,f}$, for $f = 1, \ldots, F$, most likely indicate large values of s_i^{CV} and $s_{i,f}^{PV}$, for $f = 1, \ldots, F$, which in turn should lead to a higher final price w_i. There is therefore potentially a positive relationship between the error term and the wage in the selected sample, so we suspect that $E[u_{i,f_i^w}|w_i, x_{i,f_i^w}, x_{i,f_i^w}^{inc}]$ increases with w_i. A strong positive OLS estimate of the wage effect could thus merely be an artefact of this endogeneity bias.[16]

Our Monte Carlo simulations not only confirm our conjectures regarding the direction of the different biases, but also show that the naive OLS estimates tend to be far away from the true parameter values. This is also consistent with our empirical results.

Our econometric strategy consists in adding control functions in the performance equation. A control function approach seems particularly well adapted to an environment where the matching process results from an auction: a bidding model can then provide an explicit micro-foundation for the control functions. In a nutshell the approach works as follows. Let $\mathscr{I}$ denote the set of variables observed by the econometrician right at the end of the auction sequence. This set includes in particular the variables $x_{i,f}$, W_i^r, and $z_{i,f}$, for all i and f, and w_i for workers i actually sold, but excludes the performance measures. Here $z_{i,f}$ denotes a vector of variables observed by all bidders right before the auction for worker i starts. It includes covariates such as the order of sale of i in the auction sequence (whether this worker comes up for sale first, or second, etc...), the amount of money spent by f in the preceding auctions, the characteristics

[14]For notational simplicity we omit that the expectation of u_{i,f_i^w} is also conditional on f_i^w being the auction winner and conditional on $w_i \geq W_i^r$.

[15]Under $\tau = 0$ and $\beta_x = (0, \cdots, 0)$ and if β_f does not depend on f, the OLS estimator of $\beta^{inc} > 0$ (resp. $\beta^{inc} < 0$) is downward (upward) biased if $E[u_{i,f_i^w}|w_i, x_{i,f_i^w}, x_{i,f_i^w}^{inc}]$ decreases with $\beta^{inc} \cdot x_{i,f_i^w}^{inc}$.

[16]Under $\beta_x = (0, \cdots, 0)$ and $\beta^{inc} = 0$ and if β_f does not depend on f, the OLS estimator of τ is upward biased if $E[u_{i,f_i^w}|w_i, x_{i,f_i^w}, x_{i,f_i^w}^{inc}]$ increases with w_i.

of the previous workers purchased by this firm, and possibly information on the purchases of its competitors. These variables are referred to as "the auction variables". Relying on techniques from the structural econometrics of auctions, we model and estimate the expectations $E[s^{PV}_{i,f^w_i}|\mathscr{I}]$ and $E[s^{CV}_i|\mathscr{I}]$ up to a set of parameters, and then add the corresponding estimates in the performance equation. While the auction variables z are not mandatory from a theoretical point of view, they will be helpful from an empirical perspective by playing the role of exogenous shifters in the control function terms, thereby reducing the collinearity between these control terms and the other variables appearing in the performance equation. As a prerequisite to deriving the precise form of the control functions, we now turn to the description of the auction model and the corresponding equilibrium.

3 The auction model

Here we develop an auction model in which firms compete with each other to buy the services of workers who are valued, among other things, according to their expected performance. Section 3.1 defines the auction rules. Section 3.2 defines the payoff derived by each firm from hiring a given worker as a function of the observable covariates, the full set of signals and the wage. Section 3.3 develops the equilibrium analysis.

3.1 Auction rules

Workers are sold sequentially through English auctions with public reserve prices. We formalize the game with the English button auction model (Milgrom and Weber (1982)). For worker i, the incumbent (if any) first decides whether to enter the auction at the reserve price W^r_i. Given this observable entry decision, non-incumbents then choose whether to enter the auction.[17] There are now three possibilities. First, if there are no entrants at all, the auction stops and the worker is not employed. Second, if there is a single entrant, the auction also stops but the worker is employed by this entrant and at the reserve price. Finally, if there are multiple entrants, the auction clock starts ticking at W^r_i, and moves up the price continuously. As the price goes up, entrants have to decide constantly whether to remain active in the auction, or to exit irrevocably. The clock stops when there is only one remaining active bidder. This bidder becomes the winner, and the wage w paid to the worker corresponds to the auction termination/final price.[18] Furthermore, at any price reached by the clock, we assume that

[17]The simultaneous or sequential nature of entry decisions among non-incumbents does not matter. Our modelling choice where the informed bidder enters first is motivated by the consideration that non-participation decisions are typically not irrevocable in practice. In our equilibrium analysis, we can easily check that no firm would benefit from re-entry. By contrast, if we consider alternatively that the incumbent and the non-incumbents move simultaneously at the entry stage, then the equilibrium would be such that some bidders would regret their entry decision after seeing the incumbent's decision.

[18]The tie-breaking rule (stipulating what happens when several bidders exit simultaneously at the termination price) is left unspecified since it does not play any role in our equilibrium analysis.

non-incumbents observe perfectly whether the incumbent is still active or not.

3.2 Bidders' preferences and cutoff bid-signals

We assume that firms are risk-neutral and hence maximize their expected payoff. We furthermore assume that the payoff derived by firm f from losing the auction for worker i is normalized to zero,[19] while the payoff from winning the auction at wage w is given by

$$V_{i,f} - w \equiv \overline{V}_f(x_{i,f}, z_{i,f}) \cdot e^{\lambda \cdot [s^{id}_{i,f} + s^{co}_i]} \cdot E_{\epsilon_{i,f}}[e^{\lambda \cdot y_{i,f}}] - w, \tag{2}$$

where $\overline{V}_f(.,.)$ is a strictly positive function of the firm-worker characteristics $x_{i,f}$ and the auction variables $z_{i,f}$, λ is a strictly positive parameter that captures to what extent performance matters in valuing workers, and $s^{id}_{i,f} \in \mathbb{R}$ (for $f = 1, \ldots, F$) and $s^{co}_i \in \mathbb{R}$ are additional single-dimensional signals that are discussed below.

The term $V_{i,f}$ corresponds to firm f's valuation for employing worker i. Since the wage is allowed to influence performance, the valuation itself depends on wage as well, and hence firms' preferences do no longer satisfy the "quasi-linearity assumption" that prevails in standard auction models. If in addition we have $\tau > 0$, our labor auction model implies that firms do not necessarily wish to pay the smallest possible wage to a worker. The firm-worker characteristics $x_{i,f}$ impact the valuation through two different channels: an indirect effect on firm f's payoff through the term $E_{\epsilon_{i,f}}[e^{\lambda \cdot y_{i,f}}]$, and a direct effect through the function $\overline{V}_f$. The direct effect captures the possibility that f values firm-worker characteristics intrinsically.[20]

The signals $s^{id}_{i,f}$, $f = 1, \ldots, F$ and s^{co}_i, which are normalized such that their mean is zero, capture unobserved characteristics that do not affect our performance measure $y_{i,f}$ but are are valued by firms for other reasons. The information-structure concerning these additional signals is analogous to the one adopted for the signals related to performance: for any given worker i, each firm f is assumed to receive the private signal $s^{id}_{i,f}$ and the incumbent (if any) receives also the private signal s^{co}_i. Like the signal s^{CV}_i, s^{co}_i reflects insider information held by the incumbent thanks to the previous relationship with the worker. Finally, we let $s_{i,f}$ denote the vector of firm f's private signals, i.e., $s_{i,f^{inc}_i} = (s^{id}_{i,f^{inc}_i}, s^{PV}_{i,f^{inc}_i}, s^{co}_i, s^{CV}_i)$ for the incumbent and $s_{i,f} = (s^{id}_{i,f}, s^{PV}_{i,f})$ if $f \neq f^{inc}_i$. Throughout our analysis we make the following two independence assumptions.

A1: i) The vector $s_i := ((s^{id}_{i,f}, s^{PV}_{i,f})_{f=1,\ldots,F}, s^{co}_i, s^{CV}_i)$ is i.i.d. across i. ii) For each i, the signals $s_{i,f}$ are distributed independently across f.

The first assumption stated in A1 is standard in the empirical auction literature, and ensures

[19] Setting the payoff of losing bidders at zero is not an innocuous normalization: it means that we exclude allocative externalities à la Jehiel and Moldovanu (1996) where losing bidders care about the identity of the winning bidder. From a technical perspective, this would open the door to equilibrium multiplicity. We also exclude financial externalities à la Maasland and Onderstal (2007) where losing bidders care about the price paid by their competitors (e.g., to exhaust their budget).

[20] Firms may, for instance, put higher value on employing younger workers (independently of whether younger workers perform differently relatively to older ones) because they represent a long-term investment.

that our estimators have the usual asymptotic properties. The second assumption is required for the auction model to be non-parametrically identified. It also matters for our equilibrium characterization. Note that ii) does not impose $s_{i,f}$ to be identically distributed across f, nor does it restrict the various signals belonging to $s_{i,f}$ to be independent.

Our analysis deals with the sequential/dynamic aspect of the game through a reduced form approach by including the auction variables $z_{i,f}$ into the specification of $V_{i,f}$. This vector can include e.g. the amount of money spent by f in the auctions preceding the auction for worker i, the number and characteristics of workers bought prior to i (capturing possible substitutabilities or complementarities between different types of employees), and so forth. Instead of defining firms' preferences for any bundle of workers, we define firms' preferences for any given worker and any given auction covariates. This is a reduced-form approach to capture continuation values as developed by Jofre-Bonet and Pesendorfer (2003) in a dynamic game perspective. In empirical studies on procurement data, it is now standard to use backlog variables as covariates. From a strategic perspective, it allows us to analyze bidding behavior of firms as if each worker is sold in isolation.

Let us now introduce the following renormalization of the signals: $bs_i^k := \frac{\lambda}{1-\lambda\tau} \cdot s_i^k$ for $k = co, CV$ and $bs_{i,f}^k := \frac{\lambda}{1-\lambda\tau} \cdot s_{i,f}^k$ for $k = id, PV$. These normalized signals are relevant for the analysis of the bidding stage and are therefore called the "bid-signals". We furthermore let $bs_{i,f}$ be the sum of the bid-signals observed by bidder f: $bs_{i,f_i^{inc}} = bs_{i,f_i^{inc}}^{id} + bs_{f^{inc}}^{PV} + bs_i^{co} + bs_i^{CV}$ and $bs_{i,f} = bs_{i,f}^{id} + bs_{i,f}^{PV}$ for $f \neq f^{inc}$. The CDF of $bs_{i,f_i^{inc}}$ is denoted $G_{f^{inc}}$. We also let G_f^{PV} (resp. G^{CV}) denote the CDF of $bs_{i,f}^{id} + bs_{i,f}^{PV}$ (resp. $bs_i^{co} + bs_i^{CV}$). All these CDFs are assumed to be atomless and, to simplify the exposition, to have full support on $\mathbb{R}$. The associated density functions are denoted by the lowercase letter g, e.g., g_f^{PV} corresponds to the PDF of $bs_{i,f}^{id} + bs_{i,f}^{PV}$. As detailed below, thanks to Assumption A1 and the structure of bidders' preferences as defined in (2), the equilibrium strategy of bidder f in the auction for worker i depends on the signals $\{s_{i',f}\}_{i'=1,\ldots,N}$ only through the single-dimensional aggregate bid-signal $bs_{i,f}$.

For any worker i and wage $w > 0$, and conditional on the full set of aggregate bid-signals of firm f's competitors, there is a unique aggregate bid-signal cutoff in $\mathbb{R}$ that makes firm f indifferent between hiring worker i or not. Let us characterize this cutoff first for the incumbent and then for the non-incumbents. For the incumbent this cutoff does not depend on the signals of its competitors and is then denoted $\widehat{bs}_{i,f_i^{inc}}(w)$. Defining $A_{i,f} := \frac{1}{1-\lambda\tau} \cdot \log(\overline{V}_f(x_{i,f}, z_{i,f})) + \frac{1}{1-\lambda\tau} \cdot \log[E[e^{\lambda \cdot \epsilon_{i,f}}]] + \frac{\lambda}{1-\lambda\tau} \cdot [\beta_f + \beta_x \cdot x_{i,f} + \beta^{inc} \cdot x_{i,f}^{inc}]$, we have from (2) that

$$\widehat{bs}_{i,f_i^{inc}}(w) = \log(w) - A_{i,f_i^{inc}}. \tag{3}$$

Furthermore, below (resp. above) this cutoff the incumbent's payoff is positive (resp. negative). The function $w \to \widehat{bs}_{f_i^{inc}}(w)$ is also an increasing bijection from $(0, +\infty)$ to $\mathbb{R}$ and we let then $[\widehat{bs}_{f_i^{inc}}]^{-1}(.)$ be the associated inverse function.

Non-incumbents do not observe the bid-signals bs_i^{co} and bs_i^{CV} that enter their payoff function

(2). Their aggregate bid-signal cutoffs depend thus on their competitors' aggregate bid-signals through the belief they induce on these two bid-signals and more precisely on the sum $bs_i^{co} + bs_i^{CV}$. Given A1, the distribution of $bs_i^{co}+bs_i^{CV}$ conditional of the vector of aggregate bid-signals $bs_{i',f}$ (for $i' = 1,\dots,N$ and $f = 1,\dots,F$) depends solely on $bs_{i,f_i^{inc}}$. With this is mind and letting $\theta := \lambda \cdot \tau$, we can now define the cutoff bid-signals of non-incumbents by introducing the function $U_f : \mathbb{R} \to \mathbb{R}$ given by[21]

$$U_f(x) := \frac{1}{(1-\theta)} \cdot \log(E[e^{(1-\theta)\cdot(bs_i^{co}+bs_i^{CV})}|bs_{i,f}^{id} + bs_{i,f}^{PV} + bs_i^{co} + bs_i^{CV} = x]). \tag{4}$$

For worker i, when firm $f \neq f_i^{inc}$ believes that the incumbent's aggregate bid-signal is x (and independently of its beliefs about the other competitors' bid-signals), there is, for any w, a unique aggregate bid-signal cutoff $\widehat{bs}_{i,f}(w,x)$ such that the payoff of firm f equals zero, and below (resp. above) which the payoff is positive (resp. negative). From (2) we have that

$$\widehat{bs}_{i,f}(w,x) = \log(w) - A_{i,f} - U_{f_i^{inc}}(x). \tag{5}$$

Compared to the incumbent, the cutoff signal value of each non-incumbent is shifted upwards or downward by $U_{f_i^{inc}}(x)$: this term should be interpreted as a "signal shifter" whose magnitude depends in particular on f's expectation of $bs_i^{co} + bs_i^{CV}$.

We make the following additional assumptions throughout our analysis:

A2: $\theta \equiv \lambda \cdot \tau < 1$.

A3: for any firm f, U_f is differentiable with $0 < U_f'(x) < 1$ for any $x \in \mathbb{R}$.

A4: In the auction for worker i, the distribution $G_{f_i^{inc}}$, the function $U_{f_i^{inc}}(.)$, the scalar $A_{i,f_i^{inc}}$ and the parameter θ are assumed to be common knowledge across bidders.

We do not impose τ (or equivalently θ) to be positive. Both the literature on efficiency wages (Shapiro and Stiglitz (1984)) and fairness/reciprocity in labor markets (Akerlof and Yellen (1990)) suggest, however, that we should expect τ to be positive. A2 imposes that θ is not too large in order to avoid the unrealistic setting in which bidders are prepared to pay any wage for a given worker. A3 is similar to the kind of assumptions frequently made in the literature on auctions with interdependent values (see Krishna (2002)), and implies that, for any value of w and any worker i, the marginal effect of the incumbent's bid-signal $bs_{i,f_i^{inc}}$ on $\log V_{i,f^{inc}}$ is larger than the marginal effect on $\log[E(V_{i,f}|bs_{i,f^{inc}})]$ for $f \neq f_i^{inc}$.

3.3 Equilibrium analysis

Let us first consider the situation where there is no incumbent. The sum of bid-signals $bs_i^{co} + bs_i^{CV}$ is then still unknown but, since there is no incumbent, each bidder f can only take the unconditional expectation with respect to this sum. The analogue of the signal shifter $U_f(x)$

[21] Given A1 i), the function U_f does not depend on i.

is $\frac{1}{(1-\theta)} \cdot \log(E[e^{(1-\theta)\cdot(bs_i^{co}+bs_i^{CV})}])$. The cutoff bid-signal that makes firm f indifferent between hiring or not worker i at wage w, denoted by $\widehat{bs}_{i,f}(w)$, is now given by

$$\widehat{bs}_{i,f}(w) = \log(w) - A_{i,f} - \frac{1}{(1-\theta)} \cdot \log(E[e^{(1-\theta)\cdot(bs_i^{co}+bs_i^{CV})}]). \tag{6}$$

Note that the function $\widehat{bs}_{i,f}(.)$, for $f \neq f_i^{inc}$, is strictly increasing, and that each firm f has all required information to determine this function.[22] The (weakly) dominant strategy of each firm f consists in not entering the auction for worker i if $bs_{i,f} < \widehat{bs}_{i,f}(W^r)$, and in remaining active until $[\widehat{bs}_{i,f}]^{-1}(bs_{i,f})$ otherwise. If the wage is assumed to have no effect on performance, i.e., $\tau = 0$ and hence $\theta = 0$, the model becomes a standard private value (PV) model where in the auction for worker i, it is a dominant strategy for each firm f to bid/remain active until its valuation $V_{i,f} = e^{[A_{i,f}+bs_{i,f}]} \cdot E[e^{(bs_i^{co}+bs_i^{CV})}]$.

Next we consider the situation where there is an incumbent. For notational simplicity, the index i is suppressed in the remainder of this section, and the equilibrium is derived for a given vector $\{A_f\}_{f=1,\ldots F}$. Since the incumbent's payoff only depends on his own aggregate bid-signal, the equilibrium strategy of the incumbent is always as in a PV model. If his aggregate bid-signal is below $\widehat{bs}_{f^{inc}}(W^r)$, it is a dominant strategy not to participate in the auction; otherwise, for any given price $p \geq W^r$, it is a dominant strategy to remain active (resp. to exit) if his bid-signal is above (resp. below) $\widehat{bs}_{f^{inc}}(p)$.

For the non-incumbents, the bidding incentives depend on their belief about the incumbent's aggregate bid-signal and whether the incumbent is still active in the auction. In order to determine whether a bidder $f \neq f^{inc}$ prefers to enter or not the auction, and to remain active or to drop-out at a given position of the auction clock $p \geq W^r$, and this given the bidding history up to p, we distinguish three cases.

Case A: the incumbent has decided not to enter the auction. This indicates that the incumbent's aggregate bid-signal $bs_{f^{inc}}$ is below $\widehat{bs}_{f^{inc}}(W^r)$. From Bayesian updating, the non-incumbents can then infer that $bs_{f^{inc}}$ is distributed on $(-\infty, \widehat{bs}_{f^{inc}}(W^r)]$ according to the distribution $x \to \frac{G_{f^{inc}}(x)}{G_{f^{inc}}(\widehat{bs}_{f^{inc}}(W^r))}$. Then for any $p \geq W^r$, we let $\widehat{bs}_f^A(p) := \log(p) - A_f - U_{f^{inc}}^A$ where $U_{f^{inc}}^A = \frac{1}{(1-\theta)} \cdot \log(E[e^{(1-\theta)\cdot(bs^{co}+bs^{CV})}|bs_{f^{inc}} \leq \widehat{bs}_{f^{inc}}(W^r)]) = \frac{1}{1-\theta} \cdot \log(\int_{-\infty}^{\widehat{bs}_{f^{inc}}(W^r)} e^{(1-\theta)U_{f^{inc}}(x)} \cdot \frac{dG_{f^{inc}}(x)}{G_{f^{inc}}(\widehat{bs}_{f^{inc}}(W^r))})$. Here $\widehat{bs}_f^A(p)$ corresponds to the value of the aggregate bid-signal that makes bidder f indifferent between winning the auction or not at price p, given that the incumbent has not entered the auction.

Case B: the incumbent has entered the auction and dropped out at price $p^* \in [W^r, p)$. Non-incumbents can then infer that the incumbent's aggregate bid-signal is exactly equal to $\widehat{bs}_{f^{inc}}(p^*)$. We define $\widehat{bs}_f^B(p, p^*) := \widehat{bs}_f(p, \widehat{bs}_{f^{inc}}(p^*))$, which is the value of the aggregate bid-

[22] We can rewrite $\frac{1}{(1-\theta)} \cdot \log(E[e^{(1-\theta)\cdot(bs_i^{co}+bs_i^{CV})}]) = \frac{1}{(1-\theta)} \cdot \log(\int_{-\infty}^{+\infty} e^{(1-\theta)\cdot U_{f^{inc}}(x)} dG_{f^{inc}}(x))$ for any possible f^{inc}. Given A4, θ, $G_{f^{inc}}(.)$, and $U_{f^{inc}}(.)$ are known, so (6) can indeed be determined by each firm f.

signal that makes bidder f indifferent between winning the auction or not at p, given that the incumbent has dropped out when the clock reached p^*.

Case C: the incumbent has entered the auction and is still active at price p. We let $\widehat{bs}_f^C(p) := \widehat{bs}_f(p, \widehat{bs}_{f^{inc}}(p))$, which corresponds to the aggregate bid-signal that makes bidder f indifferent between winning the auction or not at p, *as if* he knew that the incumbent would instantly exit the auction at p, which would reveal that the incumbent's aggregate bid-signal is equal to $\widehat{bs}_{f^{inc}}(p)$. Thanks to A3, $\widehat{bs}_f^C(p)$ is strictly increasing in p on $[W^r, +\infty)$.[23]

The equilibrium corresponding to the situation where an incumbent is present is given in Proposition 3.1.

Proposition 3.1. *There is a unique perfect Bayesian equilibrium in weakly undominated strategies: the strategy of the incumbent consists in not entering the auction if $bs_{f^{inc}}$ is below $\widehat{bs}_{f^{inc}}(W^r)$, and in remaining active until $[\widehat{bs}_{f^{inc}}]^{-1}(bs_{f^{inc}})$ otherwise. The strategy of non-incumbents depends on the information revealed about the incumbent's bidding behavior:*

- *Participation decision of non-incumbents: If the incumbent has not entered the auction, then the non-incumbent firm f should enter only if $bs_f \geq \widehat{bs}_f^A(W^r)$. If the incumbent has entered, then f should enter only if $bs_f \geq \widehat{bs}_f(W^r, \widehat{bs}_{f^{inc}}(W^r))$.*

- *Dropout decision of non-incumbents: Suppose the non-incumbent firm f has entered the auction, and the auction clock has reached p. If the incumbent has not entered (resp. dropped out at $p^* < p$), then f should exit instantly if $bs_f < \widehat{bs}_f^A(p)$ (resp. $bs_f < \widehat{bs}_f^B(p, p^*)$), and remain active otherwise. If the incumbent is still active at p, then f should exit instantly if $bs_f < \widehat{bs}_f^C(p)$, and remain active otherwise.*

The "inverse bidding functions" $\widehat{bs}_{f^{inc}}(.)$, $\widehat{bs}_f^A(.)$, $\widehat{bs}_f^B(.; p^*)$ (for any $p^* \geq W^r$) and $\widehat{bs}_f^C(.)$ are increasing. Given A4, each bidder f disposes of all necessary information to compute these functions.

The equilibrium strategy of each non-incumbent depends on whether the incumbent is active or not. If the incumbent is no longer active at p, then all remaining bidders share the same belief about the common value component and bidding incentives are similar to the ones in a PV model from a strategic perspective: it is a dominant strategy for non-incumbents to remain active at p if and only if their aggregate bid-signal is above the cutoff value signal that makes them indifferent between winning the auction or not at p conditional on their belief on the distribution of $bs^{co} + bs^{CV}$. Uniqueness of the equilibrium strategies is obtained exactly as in Bikhchandani et al. (2002). If, however, the incumbent is still active at p, our auction model is fundamentally an interdependent value model. What remains to be checked is that the strategy consisting in remaining active until the price $[\widehat{bs}_f^C]^{-1}(bs_f)$ is the unique optimal strategy in case C for the non-incumbent f. The formal proof can be found in Appendix A.1.

[23]Formally, this is so because $\frac{d\widehat{bs}_f^C(p)}{dp} = \frac{1}{p} \cdot [1 - U'_{f^{inc}}(\widehat{bs}_{f^{inc}}(p))] > 0$.

The analysis is similar to Milgrom and Weber (1982): non-incumbents should bid as if they are in the worst-scenario wherein the incumbent has the lowest possible signal given the bidding history.

The main take-away from this section, and which will be used throughout our econometric analysis, is that the equilibrium is unique[24] and involves strategies that are increasing in the aggregate bid-signals. Those are the key properties that need to be derived if one wants to adapt easily our approach to other auction mechanisms. Note that such properties are verified in a large class of standard auctions under pure private values (Reny (2011)).

4 The econometric methodology

This section is devoted to the econometric aspects of the paper. Section 4.1 establishes conditions under which the primitives of the auction model and the parameters of the performance equation are identified. In Section 4.2 we show how to compute in general the control terms that we plug into the performance equation to control for the various sources of endogeneity. Furthermore, we provide details on the expressions obtained under a Gaussian structure. Section 4.3 presents briefly our estimation method.

4.1 Identification

Identification is obtained in two steps: 1) From the bidding data we identify the distribution of the bid-signals, i.e., the CDFs G^{CV} and G_f^{PV} for all f, and also the parameter θ and the coefficients $A_{i,f}$ for all i and f. This is a non-parametric identification result in the sense that the bid-signal distributions are left unspecified. 2) From the first step, we can identify the various equilibrium bidding functions $\widehat{bs}$ defined in the previous section. This enables us then to identify the control terms which are added to equation (1). Finally, the augmented performance equation enables us to identify the coefficients τ, β^{inc} and β_x.

Identification requires a few additional assumptions. Before introducing them we let $S_i \subseteq \{1,\dots,F\}$ denote the publicly observable subset of firms that are potential participants in the auction for worker i.[25] In our equilibrium analysis in Section 3, we had implicitly $S_i = \{1,\dots,F\}$, but it is straightforward to adapt our equilibrium characterization for any set S_i. According to this perspective, we say that there is no incumbent if $f_i^{inc} \notin S_i$.

A5: The information set $\mathscr{I}$ contains the variables $z_{i,f}$, $x_{i,f}$, S_i and W_i^r for all i and f, the

[24]Note that this stands in contrast with the symmetric model of Milgrom and Weber (1982) which involves a continuum of perfect Bayesian equilibria in weakly undominated strategies as shown by Bikhchandani et al. (2002). However, from an econometric perspective, what we need is that the observed outcome of the game is the same across all equilibria.

[25]As is common in the auction literature, a potential participant is defined as a firm that is allowed to bid in the auction. In a procurement setup, for instance, S_i would correspond to the set of firms having passed a qualification phase in order to be allowed to submit an eligible bid.

identity of the incumbent f_i^{inc} for all workers i, and also w_i and f_i^w for all workers i that are actually sold.

A6: The signals (s_i^{co}, s_i^{CV}) are drawn independently of the signals $\{(s_{i,f}^{id}, s_{i,f}^{PV})\}_{f=1,\ldots,F}$.

A7: $\log(\overline{V}_f(x_f, z_f)) = \alpha_f + \alpha_x \cdot x_f + \alpha^{inc} \cdot x_f^{inc} + \alpha_z \cdot z_f$. For any given bidder f, there exists a set of potential participants S with $f \in S$ and $|S| \geq 2$ such that there is a positive probability that $S_i = S$ and S_i involves no incumbent while the (random) vector $(1, x_f, z_f)$ is of full rank in the corresponding subsample.

A8: The set of potential participants S_i and the incumbent's identity $f_i^{inc} \in \{0, 1, \ldots, F\}$ are drawn independently of the bid-signals, and we have: i) For each i there is a positive probability that S_i involves an incumbent. ii) In the sub-sample of auctions i with an incumbent, there is a positive probability that S_i contains exactly two bidders.

A5 determines a set of variables that is assumed to be observed by the econometrician. Note that it contains the explanatory variables appearing in performance equation (1) not only for the winning firm but also for losing bidders. Note also that we do not assume that the econometrician observes the drop-out decisions of all bidders. In particular we do not require the losing bid of the incumbent to be observed. This introduces a form of unobserved heterogeneity which complicates the analysis. A6 is a crucial independence assumption which enables us to use a deconvolution argument: once the distributions G_f and G_f^{PV} are identified for a given firm f, then we can recover the distribution G^{CV}, which is a key step to identify the functions $U_f(.)$, $f = 1, \ldots, F$, and so the bidding cutoffs associated to observed behavior when the auction involves an incumbent. The main role of the parametric specification in A7 is to simplify the presentation and, as explained later, it can be weakened. A8 is a kind of exclusion restriction:[26] in particular, the bid-signal distribution is the same in the samples with and without incumbents, which coupled with A6 enables us to disentangle the distribution of the private value bid-signal $bs_{i,f_i^{inc}}^{id} + bs_{i,f_i^{inc}}^{PV}$ and the distribution of the common value bid-signal $bs_i^{co} + bs_i^{CV}$ from the incumbent's bidding behavior. A8 ii) is a condition which simplifies the proof and can be dropped provided there is sufficient variation in the covariates as discussed in Section 5.

4.1.1 Identification of the auction model

To simplify the proof we assume in this section that the reserve price is always zero: $W_i^r = 0$ for all i. Section 5 discusses how identification can be achieved when this assumption is relaxed. In order to identify the various elements of the auction model, the proof consists in successively using different types of sub-samples (made up of specific sets of potential participants). All details of this proof are relegated to Appendix A.2. The guiding lines are sketched below.

We first use the sub-sample of auctions i for which the set of potential participants does not include an incumbent (i.e., $f_i^{inc} \notin S_i$). Given A7, the cutoff bid-signal (6) of firm $f \in S_i$

[26]See Footnote 6 for the use of similar exclusion restrictions in the literature.

associated to wage w_i can be rewritten as

$$\widehat{bs}_{i,f}(w_i) = \log(w_i) - \beta_f^* - \beta_x^* \cdot x_{i,f} - \beta_z^* \cdot z_{i,f} \equiv \log(w_i) - A_{i,f}^*$$

where $\beta_f^* = \frac{1}{1-\theta} \cdot \left[\lambda\beta_f + \alpha_f + \log[E[e^{\lambda\epsilon_{i,f}}]] + \log(\int_{-\infty}^{+\infty} e^{(1-\theta)x} dG^{CV}(x))\right]$, $\beta_x^* = \frac{1}{1-\theta} \cdot \left[\lambda\beta_x + \alpha_x\right]$ and $\beta_z^* = \frac{1}{1-\theta} \cdot \alpha_z$.

We know from Athey and Haile (2002) that the independent asymmetric private values model is non-parametrically identified under the English button auction (when there are at least two bidders). Identification hinges in particular on the following result: if a finite set of variables are drawn independently on $\mathbb{R}$, then the underlying (atomless) CDFs of these variables can be recovered from the observation of the second-order statistic coupled with the identity associated to the first-order statistic. This result can be transposed to our setting. Indeed, since bid-signals are distributed independently across bidders (A1 ii)), drop-out bids are also distributed independently across bidders conditional on any set of covariates $\{x_f, z_f\}_{f=1,\dots,F}$. This implies then that G_f^{PV} and A_f^* are identified for each f and each (x_f, z_f) in the support of the set of covariates. Given A7, this furthermore implies that β_f^* is identified for each f, as well as the vector of coefficients (β_x^*, β_z^*).

Next we exploit a sub-sample of auctions i for which S_i is composed of exactly two bidders, f and f', and where one of them, say f, is the incumbent. Then the drop-out prices are independently distributed (conditional on the observable covariates) since their bid-signals are independently distributed.[27] Again using Athey and Haile (2002), we obtain then that the distributions of the drop-out price of both bidders are identified. For the incumbent f, this means that G_f and A_f are identified. Since G_f^{PV} has been identified from the sub-sample without incumbents, a standard deconvolution argument now implies, thanks to A6, that G^{CV} is identified as well.[28] Identification of A_f for a given set of covariates (x_f, z_f) enables us to identify the constant $\beta^{inc,*} := \frac{1}{1-\theta} \cdot \left[\lambda\beta^{inc} + \alpha^{inc}\right]$ as a function of θ (and other entities that are already identified). Formally we have:

$$\beta^{inc,*} = A_{i,f} - \beta_f^* - \beta_x^* \cdot x_{i,f} - \beta_z^* \cdot z_{i,f} + \frac{1}{1-\theta} \log(\int_{-\infty}^{+\infty} e^{(1-\theta)x} dG^{CV}(x)). \tag{7}$$

Finally, identification of θ follows from the distribution of the dropout price of the incumbent's opponent, i.e., bidder f'. The relation between bidder f''s drop-out price p for worker i and its signal $bs_{i,f'}$ is given in equilibrium by $bs_{i,f'} = \log(p) - A_{i,f'} - U_f(\log(p) - A_{i,f})$ where the constants $A_{i,f'}$ and $A_{i,f}$ but also the CDF of the bid-signal $bs_{i,f'}$ have already been identified from our previous steps. Intuitively, the distribution of firm f''s drop-out price allows us to

[27] This is where the two bidders restriction plays a role: drop-out bids would no longer be independent with more than two bidders (as explained in the Appendix).

[28] The general expression of G^{CV} as an explicit function of G_f^{PV} and G_f relies on Fourier transformations. In the structural econometrics of auctions, Li et al. (2000) are the first to develop a nonparametric estimation procedure based on Fourier transformations.

recover U_f and then θ given that $U_f(x)$ is strictly increasing in θ for any given x (provided that $Var(bs^{co}+bs^{CV})>0$, a condition which would fail if we were in a private value environment). Our identification result regarding the primitives of the auction model is summarized in the following proposition.

Proposition 4.1. *In the English auction without reserve prices, the distributions G^{CV} and G_f^{PV}, $f=1,\dots,F$, the scalars θ, $\beta^{inc,*}$ and β_f^*, $f=1,\dots,F$, and the vectors β_x^* and β_z^* are identified under A1-A8.*

From Proposition 4.1, we have identified all the primitives that enable us to compute equilibrium strategies as a function of the bid-signals, and which are in turn sufficient to identify the terms we use as control functions (see Section 4.1.2).

Before ending this section we wish to make a few additional comments. 1) To simplify the presentation, we have assumed that β^{inc}, β_x and β_z do not depend on f. However, we could easily account for such asymmetries (but then A8 should be strengthened a bit). 2) Similarly, to keep the exposition as simple as possible, $\log(\overline{V}_f)$ has been specified as a linear function of the variables and parameters. We know from Athey and Haile (2007) that the vector $\{A_{i,f}\}_{f=1,\dots,F}$ can be identified as a non-parametric function of the observable vector of covariates $x_{i,f}$, $x_{i,f}^{inc}$, $z_{i,f}$ $(f=1,\dots,F)$. This means that the function $\overline{V}_f$ could be specified more generally, e.g. as a function of the competing bidders' covariates and of interaction terms between the x and z variables. Furthermore, provided that the full rank condition A7 is strengthened, we could also consider interaction terms between firm-worker characteristics and the incumbency indicator, thereby allowing the latter variable to affect valuations in a richer way (i.e., not solely through an additive shift). Finally, we could also include the dummy variable $\sum_{f=1}^F x_{i,f}^{inc}$ capturing whether i was previously employed in one of the firms (as we do in our empirical application). 3) To identify θ, it is actually sufficient to identify $U_f(x)$ for a given x. Only required is thus a given quantile of the drop-out price distribution of the non-incumbent f' (in the sub-sample of auctions with an incumbent). Contrary to the independent pure private value model under A5, our model is thus over-identified. 4) The bidding data alone do not allow us to identify the key parameters β^{inc} and τ: only $\beta^{inc,*}$ $(=(\lambda\beta^{inc}+\alpha^{inc})/(1-\theta))$ and θ $(=\lambda\cdot\tau)$ can be recovered. Nevertheless, since λ is assumed to be positive, estimation of θ allows us to determine the sign of τ, and whether it differs from zero. Performance data are needed to get the point estimates of β^{inc} and τ (and also β_x) as developed next.

4.1.2 Identification of the performance equation

The control function approach amounts to modeling the conditional expectation $E(u_{i,f_i^w}|\mathscr{I})$ using our auction model and bidding data. By definition of the bid-signals we have

$$E(u_{i,f_i^w}|\mathscr{I})=\frac{1-\lambda\tau}{\lambda}\cdot E[bs_{i,f_i^w}^{PV}+bs_i^{CV}|\mathscr{I}]. \tag{8}$$

The following assumption imposes a semiparametric restriction on the link between on the one hand the distributions of $bs_{i,f}^{id}$ and $bs_{i,f}^{PV}$ for each f, and on the other the distributions of bs_i^{co} and bs_i^{CV}.

A9: $E[bs_{i,f}^{PV}|bs_{i,f}^{id}+bs_{i,f}^{PV}=x]=\sum_{l=1}^{L}d_{l,f}^{PV}\cdot x^l$ and $E[bs_i^{CV}|bs_i^{co}+bs_i^{CV}=x]=\sum_{l=1}^{L}d_l^{CV}\cdot x^l$, for any x and each f, and the functions $x\to\sum_{l=1}^{L}d_{l,f}^{PV}\cdot x^l$ ($f=1,\dots,F$) and $x\to\sum_{l=1}^{L}d_l^{CV}\cdot x^l$ are nondecreasing.[29]

We do not impose any restriction on L, the order of the polynomial.[30] The monotonicity restriction is mild[31] and implies in particular that if a given non-incumbent bids more conditional on the observable covariates, then the worker is expected to perform better if hired by this firm. This restriction is not required to obtain our identification result but is nonetheless useful in our empirical application because it ensures the coefficients associated with our control functions terms to have a specific sign, which allows us in turn to rely on one-sided tests. Note however that we do not put any restriction on the sign of the correlation either between $s_{i,f}^{id}$ and $s_{i,f}^{PV}$ or between bs_i^{co} and bs_i^{CV}.

For each $l=1,\dots,L$, let us introduce the notation $\gamma_{l,f_i^w}^{PV}=d_{l,f_i^w}^{PV}\cdot\frac{1-\theta}{\lambda}$, $\gamma_l^{CV}=d_l^{CV}\cdot\frac{1-\theta}{\lambda}$ and

$$CF_i^{PV}[l]=E[(bs_{i,f_i^w}^{id}+bs_{i,f_i^w}^{PV})^l|\mathscr{I}]\ \text{ and }\ CF_i^{CV}[l]=E[(bs_i^{co}+bs_i^{CV})^l|\mathscr{I}].$$

When $L=1$, these entities are denoted CF_i^{PV}, CF_i^{CV}, $\gamma_{f_i^w}^{PV}$ and γ^{CV}.

From the law of iterated expectations, we have $E[bs_{i,f_i^w}^{PV}|\mathscr{I}]=E[E[bs_{i,f_i^w}^{PV}|\mathscr{I},bs_{i,f_i^w}^{id}+bs_{i,f_i^w}^{PV}]|\mathscr{I}]$ and $E[bs_i^{CV}|\mathscr{I}]=E[E[bs_i^{CV}|\mathscr{I},bs_i^{co}+bs_i^{CV}|\mathscr{I}]$. Given the independence assumptions A1 and A6, and the fact that the bidding strategy of each firm f depends on the vector of signals s_f only through its aggregate bid-signal $bs_{i,f}$ (Proposition 3.1), then the distribution of $bs_{i,f_i^w}^{PV}$ (resp. bs_i^{CV}) conditional on $(\mathscr{I},bs_{i,f_i^w}^{id}+bs_{i,f_i^w}^{PV})$ (resp. $(\mathscr{I},bs_i^{co}+bs_i^{CV})$) is the same as the distribution of $bs_{i,f_i^w}^{PV}$ (resp. bs_i^{CV}) conditional on $(bs_{i,f_i^w}^{id}+bs_{i,f_i^w}^{PV})$ (resp. $(bs_i^{co}+bs_i^{CV})$). Given A9, we can now rewrite (8) as

$$E(u_{i,f_i^w}|\mathscr{I})=\sum_{l=1}^{L}\left[\gamma_{l,f_i^w}^{PV}\cdot CF_i^{PV}[l]+\gamma_l^{CV}\cdot CF_i^{CV}[l]\right].\tag{9}$$

The terms $CF_i^k[l]$, for $k=PV,CV$ and $l=1,\dots,L$, are identified: from the bidding data, we have identified the distributions of the bid-signals $bs_{i,f}^{id}+bs_{i,f}^{PV}$ and $bs_i^{co}+bs_i^{CV}$ and also the function that maps those signals into the observable bidding history (thanks to Proposition 3.1). Computing the terms $CF_i^k[l]$, $k=PV,CV$, reduces thus to a Bayesian updating exercise. This exercise is slightly tedious in the English auction with interdependent values, and we therefore

[29] If the vectors (bs_f^{id},bs_f^{PV}) and (bs^{co},bs^{CV}) both follow a bivariate normal distribution with a positive correlation coefficient, then A9 is automatically satisfied and with $L=1$.

[30] Note that we could also use an alternative basis of functions instead of the polynomial basis $x,x^2,\cdots,x^L$.

[31] The restriction was implicitly made in our informal discussion (in Section 2) regarding the bias of the OLS estimators of τ and β^{inc}.

provide details of the calculations in Section 4.2, both in the general case and and under the assumption that bid-signals are normally distributed.
We obtain then the augmented performance equation

$$y_{i,f_i^w} = \beta_{f_i^w} + \beta_x \cdot x_{i,f_i^w} + \beta^{inc} \cdot x^{inc}_{i,f_i^w} + \tau \cdot \log(w_i) + \sum_{l=1}^{L} \gamma^{PV}_{l,f_i^w} \cdot CF_i^{PV}[l] + \sum_{l=1}^{L} \gamma_l^{CV} \cdot CF_i^{CV}[l] + \xi_{i,f_i^w} \quad (10)$$

where the error term $\xi_{i,f_i^w} = u_{i,f_i^w} - E[u_{i,f_i^w} | \mathscr{I}]$ is by construction uncorrelated to all regressors appearing in (10).[32]
If the vector composed of 1, x_{i,f_i^w}, $\log(w_i)$, $CF^{PV}_{i,f_i^w}[l]$ and $CF_i^{CV}[l]$, $l = 1, \ldots, L$, is of full rank then the parameters τ and β^{inc} and the vector β_x are identified from the performance data. The relevance of this "full rank condition" will be discussed in Section 4.2. If in addition we have $\tau \neq 0$, then the parameter λ is identified as well since $\lambda = \frac{\theta}{\tau}$. This in turn implies that the scalar α^{inc} (resp. the vector α_x) is identified through the relationship $\alpha^{inc} = (1-\theta) \cdot \beta^{inc,*} - \lambda \cdot \beta^{inc}$ (resp. $\alpha_x = (1-\theta) \cdot \beta_x^* - \lambda \cdot \beta_x$).

4.1.3 Private value environments

Let us briefly discuss how our methodology should be adapted if we were in a private value environment, i.e., if $Var(bs^{co} + bs^{CV}) = 0$, as excluded previously with the assumption that G^{CV} is atomless. There would essentially be two differences compared with the setting described so far. First, we would no longer be able to identify the parameter θ from the bidding data alone. However, the coefficient $\beta^{inc,*}$ would nonetheless be identified through (7) since the last term in this equation would actually shrink to null. The distributions G_f^{PV}, $f = 1, \ldots, F$, remain identifiable, and all this would enable us, as previously, to compute equilibrium strategies as a function of the bid-signals. Second, the control term $CF_i^{CV}[l]$ is equal to zero for each l such that the regression (10) is no longer appropriate and should be replaced by $y_{i,f_i^w} = \beta_{f_i^w} + \beta_x \cdot x_{i,f_i^w} + \beta^{inc} \cdot x^{inc}_{i,f_i^w} + \tau \cdot \log(w_i) + \sum_{l=1}^{L} \gamma^{PV}_{l,f_i^w} \cdot CF_i^{PV}[l] + \xi_{i,f_i^w}$. Given that θ is not identified from the first stage under private values, it implies that we would not be able to identify λ under private values and this even if $\tau \neq 0$. Formally, two different pairs (λ, θ) and (λ', θ') will lead to the same bid and performance outcomes if $\frac{\lambda}{1-\theta} = \frac{\lambda'}{1-\theta'}$. In practice, if we were close to a pure private value model, this suggests that some parameters of the model (like the coefficients θ, λ, d_l^{CV}, and $d_{l,f}^{PV}$ ($f = 1, \ldots, F$, $l = 1, \ldots, L$)) would be poorly estimated. This is consistent with unreported results from our simulation study, but, fortunately, the estimates of the key parameters β_x and τ remain satisfactory as reported in Section 6.

[32] There is no correlation because $\mathscr{I}$ includes all explanatory variables in (10), those appearing directly in the performance equation (i.e., the variables x_{i,f_i^w}, x^{inc}_{i,f_i^w} and w_i), and those appearing indirectly through the control terms $CF_i^k[l]$, $k = PV, CV$.

4.2 Control functions

In this section we detail how the control terms $CF_i^k[l]$, $l = 1, \dots L$, $k = PV, CV$, are characterized in general. To obtain the precise expressions of these functions, we need to perform Bayesian updating conditional on the information set $\mathscr{I}$ given the equilibrium behavior from Proposition 3.1. We also illustrate our computations for the case $L = 1$, assuming that bid-signals follow a symmetric Gaussian structure as defined next.

Definition 1. *Bid-signals follow a symmetric Gaussian structure if the bid-signals $bs_f^{id} + bs_f^{PV}$, $f = 1, \dots, F$, and $bs^{co} + bs^{CV}$ are distributed independently in the following way: i) $bs_f^{id} + bs_f^{PV}$ has the same distribution for each f and the corresponding distribution $G^{PV} \equiv G_f^{PV}$ is a centered normal distribution with variance σ_{PV}^2, ii) G^{CV} is a centered normal distribution with variance σ_{CV}^2.*

Let Φ denote the CDF of a standard normal distribution. Under a symmetric Gaussian structure we have then $G^k(x) = \Phi(\frac{x}{\sigma_k})$ for $k = PV, CV$, and A3 is automatically satisfied, as shown in the Appendix. We also show there that the two control functions for the case $L = 1$, i.e., CF_i^k, $k = PV, CV$, have explicit and tractable expressions which are derived using a series of well known properties on (truncated) normally distributed variables (which can be found in for example Greene (2008)).

The precise form of the control functions depends on whether an incumbent is present among the potential auction participants, and, if there is an incumbent, on the identity of the auction winner. We distinguish thus four cases: 1) There is no incumbent; 2) The winner is the incumbent; 3) The winner is not the incumbent and the auction clock stops at the reserve price (so that there is no second-highest bidder with probability one); 4) The winner is not the incumbent and the auction price is strictly above the reserve price (so that there is a second-highest bidder who can be either the incumbent or a non-incumbent).

Case 1: In the absence of an incumbent, worker i is sold to firm f_i^w at wage $w_i \geq W_i^r$ if and only if $bs_{i,f_i^w} \geq \widehat{bs}_{i,f_i^w}(w_i)$ and $\max\{\max_{f \in S_i \setminus \{f_i^w\}} \{[\widehat{bs}_{i,f}]^{-1}(bs_{i,f})\}, W_i^r\} = w_i$. Recall that the function $\widehat{bs}_{i,f}(.)$ is given by (6). Given the independence assumptions in A1, we have that conditional on $\mathscr{I}$, bs_{i,f_i^w} is distributed according to the distribution $G_{f_i^w}^{PV}$ truncated below $\widehat{bs}_{i,f_i^w}(w_i)$, i.e., it has the distribution function $\frac{G_{f_i^w}^{PV}(.) - G_{f_i^w}^{PV}(\widehat{bs}_{i,f_i^w}(w_i))}{1 - G_{f_i^w}^{PV}(\widehat{bs}_{i,f_i^w}(w_i))}$ on the interval $[\widehat{bs}_{i,f_i^w}(w), +\infty)$. Since there is no incumbent in this case, the independence assumptions A1 and A6 guarantee that there is no updating on $bs_i^{co} + bs_i^{CV}$ and the terms $CF_i^{CV}[l]$, $l = 1, \dots, L$, are constants that do not depend on i. We therefore have

$$CF_i^{PV}[l] = \int_{\widehat{bs}_{i,f_i^w}(w_i)}^{+\infty} x^l \cdot \frac{d[G_{f_i^w}^{PV}(x)]}{1 - G_{f_i^w}^{PV}(\widehat{bs}_{i,f_i^w}(w_i))} \quad \text{and} \quad CF_i^{CV}[l] = \int_{-\infty}^{+\infty} x^l \cdot d[G^{CV}(x)]. \tag{11}$$

Under the symmetric Gaussian structure and for $L = 1$, (11) reduces to

$$CF_i^{PV} = \sigma_{PV} \cdot \frac{\phi\Big(\frac{\widehat{bs}_{f_i^w}(w_i)}{\sigma_{PV}}\Big)}{1-\Phi\Big(\frac{\widehat{bs}_{f_i^w}(w_i)}{\sigma_{PV}}\Big)} = \sigma_{PV} \cdot \frac{\phi\Big(\frac{\log(w_i)-A^*_{i,f_i^w}}{\sigma_{PV}}\Big)}{1-\Phi\Big(\frac{\log(w_i)-A^*_{i,f_i^w}}{\sigma_{PV}}\Big)} \quad \text{and} \quad CF_i^{CV} = 0, \tag{12}$$

where $\phi(.)$ is the PDF of a standard normal distribution. The control function CF_i^{PV} corresponds to an inverse Mills ratio multiplied by the standard deviation σ_{PV}. It is a non-linear function of $\log(w_i)$, and the variables x_{i,f_i^w} and z_{i,f_i^w} (through A^*_{i,f_i^w}). If we limit ourselves to the sub-sample without incumbents then the vector $(\mathbf{1}, x_{i,f}, \log(w_i), CF_i^{PV})$ will be of full rank for each f, which implies that the associated coefficients β_f, β_x, τ, and γ_f^{PV} are identified exactly as in Heckman's (1979) seminal selection model. Note that the presence of z_{i,f_i^w} in the control function is not necessary, but in practice the auction variables may reduce potential colinearity problems (between on the one hand x_{i,f_i^w} and $\log(w_i)$, and on the other CF_i^{PV}), thereby facilitating the estimation of the parameters. Note also that the parameter γ^{CV} is not identifiable from the sub-sample consisting only of non-incumbents since there is no variation in CF_i^{CV} in this case.

Case 2: From Proposition 3.1 we know that the incumbent is the winner ($f_i^w = f_i^{inc}$) if and only if $bs_{i,f_i^{inc}} \geq \widehat{bs}_{i,f_i^{inc}}(w_i)$ and $\max\{\max_{f\in S_i\backslash\{f_i^{inc}\}}\{[\widehat{bs}^C_{i,f}]^{-1}(bs_{i,f})\}, W_i^r\} = w_i$, with $w_i \geq W_i^r$. Recall that for the incumbent the function $\widehat{bs}_{i,f_i^{inc}}(.)$ is given by (3). Given the independence assumptions in A1, $bs_{i,f_i^{inc}}$ conditional on $\mathscr{I}$ is distributed according to the distribution $G_{f_i^{inc}}$ truncated below $\widehat{bs}_{i,f_i^{inc}}(w_i)$, i.e., it has the distribution function $\frac{G_{f_i^{inc}}(.)-G_{f_i^{inc}}(\widehat{bs}_{i,f_i^{inc}}(w_i))}{1-G_{f_i^{inc}}(\widehat{bs}_{i,f_i^{inc}}(w_i))}$ on the interval $[\widehat{bs}_{i,f_i^{inc}}(w), +\infty)$. Furthermore, the distribution of $bs_i^{co} + bs_i^{CV}$, conditional on $\mathscr{I}$ and $bs_{i,f_i^{inc}}$, only depends on the aggregate bid-signal of the incumbent (i.e., $\mathscr{I}$ drops out). Noting that $bs^{id}_{i,f_i^{inc}} + bs^{PV}_{i,f_i^{inc}} = bs_{i,f_i^{inc}} - (bs_i^{co} + bs_i^{CV})$, we therefore have

$$CF_i^{PV}[l] = \int_{\widehat{bs}_{i,f_i^{inc}}(w_i)}^{+\infty} E_{x\sim G^{CV}(.|\tilde{x},f_i^{inc})}[(\tilde{x}-x)^l]\frac{dG_{f_i^{inc}}(\tilde{x})}{1-G_{f_i^{inc}}(\widehat{bs}_{i,f_i^{inc}}(w_i))} = \int_{\widehat{bs}_{i,f_i^{inc}}(w_i)}^{+\infty}\Big[\int_{-\infty}^{+\infty}(\tilde{x}-x)^l\cdot\frac{g^{CV}(x)g^{PV}_{f_i^{inc}}(\tilde{x}-x)}{1-G_{f_i^{inc}}(\widehat{bs}_{i,f_i^{inc}}(w_i))}dx\Big]\cdot d\tilde{x} \tag{13}$$

and

$$CF_i^{CV}[l] = \int_{\widehat{bs}_{i,f_i^{inc}}(w_i)}^{+\infty} E_{x\sim G^{CV}(.|\tilde{x},f_i^{inc})}[x^l]\frac{dG_{f_i^{inc}}(\tilde{x})}{1-G_{f_i^{inc}}(\widehat{bs}_{i,f_i^{inc}}(w_i))} = \int_{\widehat{bs}_{i,f_i^{inc}}(w_i)}^{+\infty}\Big[\int_{-\infty}^{+\infty}x^l\cdot\frac{g^{CV}(x)g^{PV}_{f_i^{inc}}(\tilde{x}-x)}{1-G_{f_i^{inc}}(\widehat{bs}_{i,f_i^{inc}}(w_i))}dx\Big]\cdot d\tilde{x}. \tag{14}$$

In these two expressions, the notation $E_{x\sim G^{CV}(.|\tilde{x},f_i^{inc})}$ stands for the expectation with respect to x (the sum of bid-signals $bs_i^{co} + bs_i^{CV}$) conditional on the identity of the incumbent, f_i^{inc}, and $\tilde{x}$ the realization of the incumbent's aggregate bid-signal $bs_{i,f_i^{inc}}$.

Under the symmetric Gaussian structure, the distribution of $bs_i^{co} + bs_i^{CV}$ (resp. $bs^{id}_{i,f_i^{inc}} + bs^{PV}_{i,f_i^{inc}}$) conditional on $bs_{i,f_i^{inc}} = x$ is a normal distribution with mean $\frac{\sigma^2_{CV}}{\sigma^2_{PV}+\sigma^2_{CV}} \cdot x$ (resp. $\frac{\sigma^2_{PV}}{\sigma^2_{PV}+\sigma^2_{CV}} \cdot x$).

Then for $L = 1$, (13) and (14) reduce to

$$CF_i^{PV} = \frac{\sigma_{PV}^2}{\sqrt{\sigma_{PV}^2 + \sigma_{CV}^2}} \cdot \frac{\phi\Big(\frac{\log(w_i) - A_{i,f_i^w}}{\sqrt{\sigma_{PV}^2 + \sigma_{CV}^2}}\Big)}{1 - \Phi\Big(\frac{\log(w_i) - A_{i,f_i^w}}{\sqrt{\sigma_{PV}^2 + \sigma_{CV}^2}}\Big)} \quad \text{and} \quad CF_i^{CV} = \frac{\sigma_{CV}^2}{\sigma_{PV}^2} \cdot CF_i^{PV}. \tag{15}$$

In (15), $CF_{i,f_i^w}^{CV}$ is a non-linear function of $\log(w_i)$, and the variables x_{i,f_i^w} and z_{i,f_i^w}. If we limit ourselves to the sample with incumbents and where the incumbent is the winner, then the coefficients β^{inc} and γ^{CV} are identified from the regression of $y_{i,f_i^w} - \beta_{f_i^w} - \beta_x \cdot x_{i,f_i^w} - \tau \cdot \log(w_i) - \gamma_f^{PV} \cdot CF_i^{PV}$ on the variables $x_{i,f_i^w}^{inc}$ and $CF_{i,f_i^w}^{CV}$ (without a constant term since the variable $x_{i,f_i^w}^{inc} = 1$ in this sub-sample). All parameters appearing on the left hand side of the regression are identified through case 1.

On the whole, cases 1 and 2 seem sufficient to ensure the full rank condition and thus to identify all parameters in the augmented performance equation (10). Again, our methodology does not require the presence of z_{i,f_i^w} for identification, but the auction variables may be useful for practical reasons.

Cases 3 and 4 are much more tedious and we therefore relegate the details to Appendix A.3. Below we give some elements to get a better understanding of the novel form of the control terms.

Case 3: Under the symmetric Gaussian structure and for $L = 1$, we get the following expression for the two control functions:

$$CF_i^{PV} = \sigma_{PV} \cdot \frac{\phi\Big(\frac{\log(W_i^r) - A_{i,f_i^w} - U_{i,f_i^{inc}}^A}{\sigma_{PV}}\Big)}{1 - \Phi\Big(\frac{\log(W_i^r) - A_{i,f_i^w} - U_{i,f_i^{inc}}^A}{\sigma_{PV}}\Big)} \quad \text{and} \quad CF_i^{CV} = -\frac{\sigma_{CV}^2}{\sqrt{\sigma_{PV}^2 + \sigma_{CV}^2}} \cdot \frac{\phi\Big(\frac{\log(W_i^r) - A_{i,f_i^{inc}}}{\sqrt{\sigma_{PV}^2 + \sigma_{CV}^2}}\Big)}{\Phi\Big(\frac{\log(W_i^r) - A_{i,f_i^{inc}}}{\sqrt{\sigma_{PV}^2 + \sigma_{CV}^2}}\Big)}, \tag{16}$$

where $U_{i,f_i^{inc}}^A$ is the third term appearing in $\widehat{bs}_{i,f_i^w}^A(W_i^r)$ as defined in Section 3.3. In Appendix A.3 we show that it equals[33]

$$U_{i,f_i^{inc}}^A = \frac{1}{1-\theta} \log\Bigg(\frac{\Phi\big(\frac{[\log(W_i^r) - A_{i,f_i^{inc}} - (1-\theta)\cdot\sigma_{CV}^2]}{\sqrt{\sigma_{PV}^2 + \sigma_{CV}^2}}\big)}{\Phi\big(\frac{[\log(W_i^r) - A_{i,f_i^{inc}}]}{\sqrt{\sigma_{PV}^2 + \sigma_{CV}^2}}\big)}\Bigg) + \frac{(1-\theta)}{2} \cdot \sigma_{CV}^2. \tag{17}$$

In cases 1 and 2, two features play in favor of the full rank condition: the control terms are non-linear in x_{i,f_i^w}, the firm-worker characteristics associated with the winner, and they depend on z_{i,f_i^w}, the auction variables of the winner which act as instruments. In case 3, we see that there

[33]This term can be interpreted as the amount by which the non-incumbents (including the auction winner) have shifted their cutoff bid-signal to account for the fact that the incumbent f_i^{inc} did not enter the auction at the reserve price W_i^r. Note that $U_{i,f_i^{inc}}^A \leq \frac{(1-\theta)}{2} \cdot \sigma_{CV}^2 = A_{i,f_i^{inc}}^* - A_{i,f_i^{inc}}$, the value of the bid-signal shifter when the reserve price goes to infinity or equivalently if the incumbent were not eligible to bid.

is another force that reduces the colinearity between the regressors and the control terms: the latter depend on the covariates of a losing bidder as well, namely the incumbent's covariates $x_{i,f_i^{inc}}$ and $z_{i,f_i^{inc}}$ (through $A_{i,f_i^{inc}}$).

A new difficulty also arises compared to cases 1 and 2: the control term CF_i^{PV} now depends on the parameter θ and not only through the term $(1-\theta)\cdot\sigma_{CV}^2$ as we can see from (17). This is potentially an issue since θ is not identified when $\sigma_{CV}=0$, and thus poorly estimated when σ_{CV} is small. However, a Taylor expansion reveals that $U^A_{i,f_i^{inc}} = -\sigma_{CV}^2\cdot\Big[\phi(\frac{[\log(W_i^r)-A_{i,f_i^{inc}}]}{\sigma_{PV}})/\Phi(\frac{[\log(W_i^r)-A_{i,f_i^{inc}}]}{\sigma_{PV}})\Big]+\frac{(1-\theta)}{2}\cdot\sigma_{CV}^2+O(\sigma_{CV}^4)$. This suggests that our methodology could work well even if σ_{CV} is small as confirmed by our simulations.

Case 4: When the final price is strictly above the reserve price and the winner is not the incumbent, the calculations of the control functions are relatively complex. The difficulty comes mainly from the fact that the bid-signals of the winner and the incumbent are no longer independently drawn conditional on $\mathcal{I}$. This is so because the bidding strategies of the non-incumbents depend on the moment when the incumbent has dropped out from the auction, and this moment is not observed by the econometrician. The calculations become simpler if the econometrician could observe the bid/wage w_i^{inc} where the incumbent has quit the auction (if the incumbent does not enter the auction, we let $w_i^{inc}=NP$). Under the symmetric Gaussian structure and for $L=1$, we obtain then the following expressions for the two control functions (denoted now by $CF_i^{PV}(w_i^{inc})$ and $CF_i^{CV}(w_i^{inc})$):

$$CF_i^{PV}(w_i^{inc})=\sigma_{PV}\cdot\frac{\phi\Big(\frac{\widehat{bs}_{f_i^w}(w_i,\widehat{bs}_{f_i^{inc}}(w_i^{inc}))}{\sigma_{PV}}\Big)}{1-\Phi\Big(\frac{\widehat{bs}_{f_i^w}(w_i,\widehat{bs}_{f_i^{inc}}(w_i^{inc}))}{\sigma_{PV}}\Big)}\ \text{ if } w_i^{inc}\neq NP,\ \text{ and } CF_i^{PV}(NP)=\sigma_{PV}\cdot\frac{\phi\Big(\frac{\widehat{bs}^A_{f_i^w}(W_i^r)}{\sigma_{PV}}\Big)}{1-\Phi\Big(\frac{\widehat{bs}^A_{f_i^w}(W_i^r)}{\sigma_{PV}}\Big)},\tag{18}$$

and

$$CF_i^{CV}(w_i^{inc})=\frac{\sigma_{CV}^2}{\sigma_{PV}^2+\sigma_{CV}^2}\cdot\Big[\log(w_i^{inc})-A_{i,f_i^{inc}}\Big]\ \text{ if } w_i^{inc}\neq NP,\ \text{ and } CF_i^{CV}(NP)=-\frac{\sigma_{CV}^2}{\sqrt{\sigma_{PV}^2+\sigma_{CV}^2}}\cdot\frac{\phi\Big(\frac{\log(W_i^r)-A_{i,f_i^{inc}}}{\sqrt{\sigma_{PV}^2+\sigma_{CV}^2}}\Big)}{\Phi\Big(\frac{\log(W_i^r)-A_{i,f_i^{inc}}}{\sqrt{\sigma_{PV}^2+\sigma_{CV}^2}}\Big)}.\tag{19}$$

If the incumbent's drop-out bid were observed, $CF_i^{CV}(w_i^{inc})$ (for $w_i^{inc}\neq NP$) can no longer be expressed as an inverse Mills ratio but is rather a linear function of the covariates of the incumbent (through $A_{i,f_i^{inc}}$), reflecting that here we are able to recover perfectly the incumbent's bid-signal so that the Bayesian updating exercise is more informative.

If the incumbent's drop-out bid were not observed, the control functions CF_i^{PV} and CF_i^{CV} would correspond to a weighted-average of the terms in (18) and (19). In the appendix we compute these weights, and show in particular that they depend on the covariates of *all* potential participants in S_i (not just of the winner and incumbent as in the previous cases), adding yet another source of variation in the control terms.

4.3 Estimation

We propose below a parametric procedure which follows our identification argument and involves thus two stages: first, the primitives of the auction model are estimated through maximum likelihood; second, the parameters of the performance equation are estimated by OLS with the estimated control functions added as controls.

The likelihood function of the auction data is given in Appendix B. The distribution functions G^{CV} and G_f^{PV}, $f = 1, \ldots, F$, are assumed to be known up to a parameter belonging to R^d for some $d \geq 1$. These parameters are estimated in the first stage together with the scalars θ and β_f^*, $f = 1, \ldots, F$, and the vectors β_x^* and β_z^*, which altogether completely characterize how firms bid as a function of their bid-signals, and this for any worker i given the set of observable covariates $x_{i,f}, z_{i,f}$, $f = 1, \ldots, F$. We will assume here that the identity of the second-highest bidder is observed by the econometrician, as is the case in our dataset.

Some remarks on the likelihood function: 1) In writing down the likelihood, one should carefully pick the appropriate density or distribution functions of the cutoff bid-signals associated to observed bidding behavior (depending on whether a worker is sold, on whether there is an incumbent, and on whether the incumbent is the winner or the second-highest bidder), but otherwise its structure is relatively simple and resembles some of the likelihoods derived in pure private value models (see Paarsch and Hong (2006)). 2) Depending on the specific parametric distribution function chosen for the bid-signals, it may occur that the support of the observables depends on the vector of parameters, or more generally that the conditional density suffers from discontinuities with respect to the parameters. This would violate the regularity conditions required to derive the usual $\sqrt{n}$-asymptotic normality of the ML estimators. In both our simulation protocol and our empirical application, the bid-signals are assumed to be normally distributed, implying that there are no discontinuities in the conditional density. We will therefore abstract from these potential additional complications (addressed by Chernozhukov and Hong (2004)). 3) The main computational burden in our likelihood function arises from calculating the signal shifters $U_f(.)$ and U_f^A, which rely on deconvolution integrals. However, under our normality assumption, these integrals simplify: in particular the function $U_f(.)$ is linear. Another computational difficulty comes from contributions where the incumbent is neither the winner nor the second-highest bidder. As detailed in Appendix B, it requires an integral over all possible realizations of the incumbent's drop-out price. In our Monte Carlo study and empirical application, this difficulty is avoided since the drop-out price of the incumbent is observed.

Our second stage consists in estimating the performance equation

$$y_{i,f_i^w} = \beta_{f_i^w} + \beta_x \cdot x_{i,f_i^w} + \beta^{inc} \cdot x_{i,f_i^w}^{inc} + \tau \cdot \log(w_i) + \sum_{l=1}^{L} \gamma_{l,f_i^w}^{PV} \cdot \widehat{CF}_i^{PV}[l] + \sum_{l=1}^{L} \gamma_l^{CV} \cdot \widehat{CF}_i^{CV}[l] + \text{error}_{i,f_i^w}, \tag{20}$$

using the performance data. Here $\widehat{CF}^{PV}_{i,f_i^w}[l]$ and $\widehat{CF}^{CV}_{i}[l]$ are the estimated control functions, i.e., the expressions one gets after replacing all unknown parameters entering the control functions by their first-stage estimates. The parameters β_f, β_x, β^{inc}, τ, $\gamma^{PV}_{l,f}$, and γ^{CV}_{l}, for $l = 1, \ldots, L$, are estimated by OLS. The resulting estimator is consistent and asymptotically normally distributed (see Wooldridge (2002)). The standard errors are obtained using a percentile bootstrap method.[34]

5 Related literature and identification issues

5.1 Links with the auction literature

Most empirical papers applying the structural econometrics of auctions assume pure private values and risk-neutrality. Our auction model and its identification from bidding data only are novel due to two main ingredients: 1) Bidders' payoff functions are no longer quasi-linear in the auction price (if $\theta \neq 0$); 2) Our model involves interdependent values with asymmetric bidders. Such ingredients have already being addressed to some extent in the literature as discussed next. We also stress that our setup involves multi-dimensional signals which is not innocuous from an econometric point of view. However, from an equilibrium analysis perspective, our auction model is equivalent to one with single-dimensional signals: as in Goeree and Offerman (2003)'s interdependent values model with bi-dimensional private signals, the bidding incentives of firm f for worker i depend on the vector of private signals $s_{i,f}$ only through the single-dimensional aggregate bid-signal $bs_{i,f}$.

In the existing literature, departures from the quasi-linear paradigm arise typically from risk aversion which is known to cause important identification problems (Guerre et al. (2009)), but also possibly from contingent payment auctions (like equity auctions) where monetary transfers depend on some ex-post realization (as in Bhattacharya et al. (2018)). In our setup, the departure from the quasi-linear paradigm is captured by the parameter θ. Once the distributions of the bid-signals of both the common value components and the private value components for a given firm f are identified, θ is identified by comparing the equilibrium bid distributions of a given bidder $f' \neq f$ in the sub-sample where there is no incumbent with the sub-sample where f is the incumbent, i.e., by comparing bids in auctions where θ drives bidding behavior with auctions where it plays no role.[35] More precisely, we argued in Section 4 that any quantile

[34] Alternatively, the standard errors could be obtained using the first-stage adjusted asymptotic variance to account for the fact that the estimated control functions have been substituted for their unknown counterparts (see Wooldridge (2002)).

[35] In a related vein, Lu and Perrigne (2008) identify jointly bidders' valuation distribution and their risk aversion -both in a non-parametric way- thanks to the observation of bidding data for both the English auction (where equilibrium strategies do not depend on bidders' risk aversion) and the first price auction (where more risk-aversed bidders bid more aggressively). In a two-stage sequential auction setup where the first auction is a first-price auction and the second an English auction, Kong (2018) identifies a model involving both risk aversion and a general form of multi-object preferences. Both Lu and Perrigne (2008) and Kong (2018) stick to pure private values with

of the bid distribution of a non-incumbent in the sample with an incumbent allows to identify θ: this suggests that much more general forms could be used for the incentive effect (instead of the linear specification $\tau \cdot \log(w)$). Extending our model in this direction, and studying the possible identification issues that may arise, is beyond the scope of this paper and is left for further research.[36]

It is well-known that models with interdependent values raise difficult issues. On the one hand, equilibrium existence can fail in the second-price auction or in the English auction as illustrated by Jackson (2009), contrary to the first-price auction where Reny and Zamir (2004) obtain a very general positive result. On the other hand, Laffont and Vuong (1996) argue that without exclusion restrictions, any bidding data generated by an interdependent value model can be rationalized by a pure private value model.

Most auction models with interdependent values impose that bidders are fully symmetric. Beyond Milgrom and Weber (1982)'s seminal model, Goeree and Offerman (2003) develop a symmetric model where each bidder receives a bi-dimensional private signal that reflects respectively a pure private and a pure common value component. By contrast, in our model, only the incumbent firm receives a signal about the common value of the worker, an ingredient we borrow from Engelbrecht-Wiggans et al. (1983)'s seminal pure common value model with an informed bidder.[37] On the whole, our model can then be viewed as an hybridization between the models of Engelbrecht-Wiggans et al. (1983) and Goeree and Offerman (2003), while allowing for general forms of asymmetries across bidders regarding private preferences.

Extending our model to allow for the information about the common value component to be dispersed (possibly asymmetrically) among multiple bidders would be of primary interest. However, it would raise important challenges both in terms of equilibrium existence and identification, at least without further restrictions on the distributions of bidders' signals. Those issues could be circumvented by imposing some parametric restrictions.[38]

5.2 Alternative approaches to identification

The identification argument we have presented does not rely on exogenous variations of the covariates (as in the bulk of the literature on the econometrics of auctions). Indeed, for any set of covariates, we can identify the primitives of interest. Nevertheless, the exclusion restriction A8 could be substituted by an assumption stating that bidder-specific covariates vary across

independent signals across bidders.

[36] Our equilibrium analysis with strictly monotonic strategies extends easily if the term $\tau \cdot \log(w)$ is replaced by some general function $\rho(w)$. Assumption A2 should then be replaced by two restrictions on this function: $\rho'(w) < \frac{1}{\lambda \cdot w}$ and $\lim_{w \to +\infty} w \cdot e^{-\lambda \cdot \rho(w)} = +\infty$.

[37] See Hendricks et al. (1994) for a generalization with a random reserve price and an application to oil tract auctions.

[38] See Hong and Shum (2003) and Heumann (2019) for interdependent values models involving a Gaussian information structure and with respectively single-dimensional and multi-dimensional signals. Relatedly, Weiergraeber and Wolf (2018) develop an empirical analysis of a generalization of Goeree and Offerman (2003)'s model but without any formal guarantees regarding identification.

auctions. As explained in Athey and Haile (2007), such variations allow identification in a much simpler manner if we are prepared to assume that those variations are exogenous. In particular, samples without an incumbent could be substituted by samples with an incumbent having unfavorable covariates ($A_{i,f_i^{inc}}$ going to minus infinity) so that the probability that the incumbent enters the auction goes to zero.

Our identification argument also abstracts from the presence of reserve prices. Binding reserve prices may prevent to identify (non-parametrically) the signal distribution over its full support. This issue can be solved by assuming that there is sufficient variations in the covariates.[39] We stress, however, that our primary goal is to control for endogeneity and adverse selection in our second stage, and for this, full-support identification is not necessarily required: if we are in a pure private value setup, for instance, then we do not care about identifying the distribution $G_f^{PV}(.)$ for the bid-signals that are below $\widehat{bs}(W_i^r)$ because we need to identify $E[bs_{i,f_i^w}^{PV}|\mathscr{I}]$ only up to a constant. Binding reserve prices could nevertheless be helpful for identification. E.g., Roberts (2013) shows that variations in the reserve price allow to deal with unobserved heterogeneity in the English auction when there is a monotonic relation between the reserve and the "quality" of the good for sale that is observed by bidders but not by the econometrician.

The English auction where only the transaction price and the winner's identity are observed is known to be the most unfavorable case for identification (Athey and Haile (2007)). Under other auction formats (like the first price auction that is commonly used in procurement) or if additional information is available in the English auction (like the identity of the second-highest bidder, the identity of the set of participating bidders, or additional information about the auction dynamic), we can not only adapt our methodology but we can actually improve it: From an identification point of view, observing bids from multiple bidders would allow us to relax either our exclusion restrictions or the independence assumption A1 ii). E.g., Li et al. (2000) show how the observation of two bids allows to deal with a structure where the correlation across bidders' valuations results from a multiplicative common shock.[40] From an estimation point of view, observing multiple bids should increase the precision of our estimates of the control terms and thus reduce the resulting finite-sample bias.

6 Simulation study

The aim of the simulation study is to show how our methodology performs using small sample sizes (we pick $N = 300$ or $N = 1,000$ auctions). The parameters of our Monte Carlo protocol have been chosen as follows. As in our empirical application, we set the total number of bidders

[39] If $A_{i,f}$ can be arbitrary large for all f and some realizations of the covariates then we would be back to the case where the reserve price is arbitrarily small and bid distributions are identified on their full support.

[40] Similarly, an important caveat of our approach is that if the econometrician does not observe some public covariates that enter the bidding equation, then it would induce implicitly a failure of the independence assumption A1 ii) which would then bias our estimations of the control terms. As in Krasnokutskaya (2011), observing two bids would allow to deal (non-parametrically) with such form of unobserved heterogeneity.

to $F = 8$. Throughout, it is assumed that the set of potential participants coincides with the full set of eight bidders. In each simulated data set, half of the sample is composed of auctions without incumbents while the other half is made up of auctions with an incumbent (picked randomly). In all simulations we fix $\lambda = 1$, while τ can take the values -0.8, -0.5, 0, 0.5 and 0.8 (these are hence also the values taken by θ). As in the illustrative example of Section 4.2, we assume that the bid-signals follow a symmetric Gaussian structure. In addition, we assume that $bs^{id}_{i,f} = bs^{co}_{i} = 0$ for all i and f, implying that $L = 1$ and $d^{CV} = d^{PV}_{f} = 1$ in A9, which in turn means that the coefficients associated with the two control functions in (10) equal $\gamma^{CV} = \gamma^{PV}_{f} = 1 - \tau$ for all f. For all simulations we set $\sigma_{PV} = 1$, while σ_{CV} take the values 0, 1 and 2 ($\sigma_{CV} = 0$ corresponds to the private value case).

We also assume that $\beta^{inc} = \alpha^{inc} = 0$, i.e., the incumbency-indicator $x^{inc}_{i,f}$ is not included in the performance and bidding equations. The vectors of covariates $x_{i,f}$ and $z_{i,f}$ are both single dimensional. In line with our empirical application, the firm-worker characteristic $x_{i,f}$ is assumed to be the same across all bidders, but the auction variable $z_{i,f}$ does vary with f. Those variables are assumed to be i.i.d. according to a centered normal distribution with variance equal to 1. We set $\beta_f = 0$ for all bidders and $\beta_x = 1$, while the noise $\epsilon_{i,f}$ is assumed to be distributed according to a centered normal distribution with variance equal to 1.[41] We also set $A_{i,f} = x_{i,f} + z_{i,f}$, which amounts to choosing the following parameter values appearing in $\overline{V}_f(x_{i,f}, z_{i,f})$: $\alpha_f = -\frac{1}{2}$, $\alpha_x = -\tau$ and $\alpha_z = 1 - \tau$. Finally, the reserve price W^r_i is set to $3 \cdot \frac{\lambda}{1-\lambda\cdot\tau} = \frac{3}{1-\tau}$.

Table 1 reports the mean estimates (over 1,000 Monte Carlo replications) of the parameters τ and β_x for three methodologies: 1) standard OLS (estimation of the performance equation without any control term); 2) the variant of our methodology sketched in Section 4.1.3 and referred to as the PV methodology;[42] 3) Our general methodology described in Section 4 and referred to as the CV methodology. For both the PV and CV methodologies we fix $L = 1$. We also report at the bottom of Table 1 the percentage of the players that are sold (when $N = 1000$): this percentage varies between 0.28 and 0.98 in our simulations.

Table 2 reports the estimated lower and upper bounds of the 95% confidence intervals for τ and β_x. We estimate these bounds using the Warp-Speed method developed by Giacomini et al. (2013). By relying on a single bootstrap sample for each replication, the Warp-Speed method reduces drastically the computational cost of our Monte Carlo experiments.[43] The Warp-Speed technique is briefly summarized in Appendix B. Also reported in Table 2 are estimates of statistical power. We calculate the power of the test of the null hypothesis $H_0 : \tau = 0$ (resp.

[41] These choices ensure that one half of the sample variation in the simulated performance measure is explained by $x_{i,f}$, about the same fraction as in our real data.

[42] The likelihood maximized at the first stage is thus the one corresponding to a standard PV model with all bidders having cutoff bid-signals of the form (6). Compared to our general specification, it is as if σ_{CV} is fixed to be zero while the parameter θ is no longer present in the likelihood.

[43] This is crucial in our study since the power and confidence interval computations for each parameter set (each column of Table 2) takes approximately one day with the Warp-speed method (i.e. with one bootstrap draw).

$H_0 : \beta_x = 0$), against the bilateral alternative, as the fraction of times zero does not belong to the 1,000 CIs for τ (resp. β_x), given that the data are generated under a particular value of τ (resp. β_x). Finally, we report the empirical coverage probabilities (i.e., the probability that the parameter's true value lies in the CI). The results in this table are only given for the CV methodology.

We see from Table 1 that, as predicted in Section 2, the OLS estimator of τ is upwards biased while the one of β_x is biased toward zero. The bias is very substantial for both parameters, especially when τ takes small values.[44] For instance, when $\tau = -0.8$ and $N = 1,000$, the mean OLS estimate of τ ranges between -0.23 ($\sigma_{CV} = 0$) and 0.17 ($\sigma_{CV} = 1$), while the mean OLS estimate of β_x (recall that its true value is 1) varies between 0.09 ($\sigma_{CV} = 1$) and 0.43 ($\sigma_{CV} = 0$). We also see that the smallest biases are always obtained when $\sigma_{CV} = 0$. For $\tau \geq 0$, the biases increase in σ_{CV}, while for the case $\tau = -0.5$ and -0.8 the strongest biases are obtained for the intermediary value $\sigma_{CV} = 1$.

The PV estimator performs well when $\sigma_{CV} = 0$. This does not come as a surprise because precisely for this value the data are generated under the PV paradigm, and the PV methodology then provides consistent estimators. We see from Table 1 that the results are not satisfactory when $\sigma_{CV} > 0$. For instance, if $\tau = -0.5$ and $\sigma_{CV} = 1$, the mean estimate of τ (resp. β_x) is around -0.21 (resp. 0.74) for both sample sizes, a relative bias of approximately 58% (resp. 26%). Nevertheless, the PV methodology produces much better results than OLS. Looking for example at the estimates of τ, the PV method reduces roughly half (resp. two thirds) of the bias of the OLS estimator when $\sigma_{CV} = 2$ (resp. $\sigma_{CV} = 1$).

Consider finally the results for the CV estimator. Its performance is comparable to the performance of the PV estimator when $\sigma_{CV} = 0$. This is quite surprising since there is no formal guarantee that our CV methodology is consistent under the PV paradigm.[45] At the same time, Table 1 shows that the CV estimator clearly outperforms the PV estimator in the case $\sigma_{CV} > 0$. Indeed, when $\sigma_{CV} = 1$, the CV estimator of τ is perfectly consistent for any value of the wage effect; the parameter β_x is also well estimated although there is a small upwards bias for negative values of τ. When $\sigma_{CV} = 2$, the CV estimator of τ (resp. β_x) is slightly downwards (resp. upwards) biased, but the the biases are much smaller than those produced by the PV estimator.

Table 2 shows that the empirical coverage rates for the CV methodology are satisfactory for both parameters: when the estimator produces an unbiased estimate, then the rate ranges between 0.94 and 0.96. As expected, lower rates are obtained, however, whenever the estimator performs less well. The lowest coverage rates are obtained for the case $N = 1000$ and $\tau = -0.5$ and $\sigma_{CV} = 2$: it equals 0.69 for τ and 0.84 for β_x.

Table 2 also shows that our tests of $H_0 : \tau = 0$ and $H_0 : \beta_x = 0$ have high power: in our

[44]This is as expected from our peculiar specification where the parameters γ_f^{PV} and γ^{CV} both decrease in τ and vanish when $\tau = 1$, so endogeneity matters more when τ is small and becomes negligible as τ goes to 1.

[45]However, this is less surprising given that the dependence of the control term CF_i^{CV} with respect to θ seems to vanish when σ_{CV} goes to zero (as formalized for case 3 in Section 4.2).

simulations with $\tau \neq 0$, the probability of correctly rejecting the null hypothesis $H_0 : \tau = 0$ fluctuates between 69 and 100% when $N = 300$, and equals at least 98% when $N = 1,000$; the probability of correctly rejecting $H_0 : \beta_x = 0$ is at least equal to 98% in the smaller samples, and always equal 100% in the larger ones. Note that the power results reported in the upper panel and under the columns headed $\tau = 0$ actually give the size of the test of $H_0 : \tau = 0$ against the two-sided alternative. As indicated in the table, the size is adequate when σ_{CV} equals 0 or 1 (it then varies between 0.04 and 0.06), but seems a bit too large when $\sigma_{CV} = 2$ (0.10 if $N = 300$ and 0.24 if $N = 1,000$). This last result can be explained by the fact that, as mentioned above, the CV estimator of the wage effect is slightly biased for this relatively high value of the standard deviation.

Finally, we see from Table 2 that the estimated lower-bounds and upper-bounds of the CIs are generally symmetrically distributed around the true values (especially those corresponding to β_x). This symmetry is weaker when $\sigma_{CV} = 2$ because the CV estimator is then slightly biased. In this case the bounds are symmetrically distributed around the estimated mean values. We observe that, as expected, the CIs are tighter for the larger samples than for the smaller ones. We also observe that the CIs for β_x become tighter as σ_{CV} gets smaller. Contrastingly, there is not a clear-cut relationship between the CIs for τ and the standard deviation σ_{CV}: for $\tau = -0.8, -0.5, 0$, the CI is tighter as σ_{CV} gets smaller, while the reverse holds for $\tau = 0.5, 0.8$. Intuitively, there are two countervailing forces when σ_{CV} gets larger: on the one hand, the selection effect becomes more important and introduces additional noise in the performance equation (formally, the variance of the residual ξ_{i,f_i^w} gets larger), reducing the precision of the estimates. On the other hand, $\widehat{CF}_i^{CV}$ becomes a less noisy estimator of the control term CF_i^{CV} (θ is no longer identified when $\sigma_{CV} = 0$ and thus CF_i^{CV} will be poorly estimated when σ_{CV} is small), thereby instead augmenting precision. Depending on the parameter being estimated (and the true value of θ), one of these forces dominates the other.

7 Data and empirical application

7.1 Tournament and player performance

The Indian Premier League (IPL) is an annual cricket tournament where teams compete by playing matches in a double round-robin format. At the end of this first stage, the four best ranked teams compete in a playoff to determine the final winner of the tournament. In our empirical analysis we focus on the 2014 IPL because it represents a year in which major player auctions were held before the tournament, whereby players were (re)allocated to teams. In that year, eight teams competed in the tournament and each team played between 14 and 16 matches depending on whether it qualified in the playoff.

A cricket match is played over a fixed time period (three hours in the IPL) between two teams consisting of 11 players who are selected from the team squads. Cricket players are categorized into four categories: batsman, bowler, wicket keeper and all-rounder. The unique feature of cricket is that, unlike most other team sports, a large component of overall team performance depends on individual specific performances. Since player skills are highly specialized, it is possible to observe a large set of individual measures of performance that are idiosyncratic and largely independent of how other team members perform. On average, we observe 115.6 performance measures per team over the 2014 tournament (note that some players in the selected squad may not perform).

From the individual player performances we construct a composite performance measure. It is derived from various, batting and bowling statistics observed for each player during the tournament. We award points for each statistic accumulated by each player across every game. The way that a player accumulates points and the construction of our composite performance measure is described in detail in Appendix B.1. We also give there additional information on the rules of the cricket game.

Data sources: We obtained performance data on all matches played in the tournament from www.espncricinfo.com. All data on player auctions, described in the following section, were manually compiled from the recordings and minutes[46] of the (publicly broadcasted) auction proceedings.

7.2 Player auctions

Beginning in 2008, once every three years, the IPL organizes auctions to (re)allocate players to teams. This centralized market is the unique opportunity for teams to hire new players. Furthermore, any player remaining unsold in the auctions does not participate in the tournament. Prior to the sale, each player is assigned a reserve price that represents the price at which bidding starts. The reserve price is broadly determined by the auctioneer based on a variety of factors, primary among them being the player's past performance.[47]

The format of sale consists of a sequential procedure whereby players are sold one after the other through a series of English auctions with public reserve prices. In each of those independent auctions, teams were invited to challenge the temporary winner by raising their paddle to indicate their willingness to buy the player at the current price plus a predetermined increment. However, our analysis abstracts from the bid increments and proceeds as if the bidding data is generated from an English button auction (see Section 3.1).[48] If a player played in one of the teams in the previous IPL season, then he is declared to be RTM-eligible, where the RTM acronym stands for "right-to-match". In an auction for a player that is not RTM-eligible, then

[46]The minutes of the live auction proceedings were obtained from ESPN-Cricinfo (http://www.espncricinfo.com/indian-premier-league-2014/content/story/718095.html).

[47]In 2014, seven different reserve prices were used from 1 up to 20 Millions of Rupees.

[48]Lamy et al. (2016) do take into account increments in their auction model but at the cost of being not fully structural.

the provisional winner (if any) was declared the final winner and the player's salary for the IPL tournament corresponds to the last bid submitted. If the player is RTM-eligible, then his team from the previous year had the option to use one of its RTM cards and match the winning offer to buy-back their player at the salary fixed by the auction.[49]

The auctioneer arranges the players into different '<u>sets</u>' by their cricketing speciality, popularity, and, to some extent, their reserve price. The sale of players proceeds according to a predetermined sequence of these sets. The composition of the sets and the sequence in which they are placed in the auction are announced ex ante. By contrast, the order in which players are auctioned within each set is determined by random draws in the format of a lottery.

The teams face a set of explicit rules with regard to both team composition and bidding behavior. These rules play an important role in determining some constraints that bidders face whilst bidding. In 2014 there were three types of rules. 1) <u>Spending cap</u>: in order to encourage a balanced competition, the organizers imposed a spending cap on the total amount that any bidder was allowed to spend in the auctions. The spending cap allocated to a bidder depends on the number of players retained by the team from its previous year's squad (the less players retained, the higher the cap). Teams were allowed to retain a maximum of five players from their previous year's squad, and the spending cap varied from 245 to 700 Millions of Rupees. 2) <u>Overseas player quota</u>: to ensure a sufficient number of native players in the tournament, the organizers imposed a maximum limit of 9 on the number of overseas (non-Indian) players in any team. 3) <u>Right-to-Match (RTM)</u> option: depending on the number of retained players, each team received from the organizers between 1 to 3 RTM cards. As presented above, a RTM card allowed a team to exert the RTM option for any player from its previous year's squad.

7.3 Descriptive statistics

A total of 317 players and 8 teams participated to the 2014 IPL auctions. Out of these 317 players, 122 received bids at or above the reserve prices and were actually sold. For all players (including unsold players) we know a number of characteristics: their nationality, their cricket speciality, and whether they are a so-called newcomer.[50] We record, for every auction, whether in the previous year the player was playing for one of the 8 teams, and, if this is the case, the identity of the player's previous team. Using the terminology previously used in the paper, the corresponding team is referred to as the incumbent. We observe how players are pooled into sets, the sequence of the sets, and the order in which players are auctioned within sets. We also observe in the data all reserve prices attached to the players, and, for 105 players among those sold, the composite performance measure defined in Appendix B.1. For each auction we

[49] The RTM option is equivalent to what is called the "right-of-first-refusal" option in the auction literature (see Bikhchandani et al. (2005) for an analysis of second-price/English auctions with such an option). Note that the team with the possibility to exert the RTM option was allowed to bid in the auction exactly as the other bidders.

[50] A newcomer is a player who has already been called by his national team. Such players are of course already experienced, but we nonetheless use the terminology "Newcomer" as this is the official designation for such players.

observe all submitted bids (i.e., all prices at which teams raised their paddles) together with the identities of the corresponding bidders, and the identity of the team who has used a RTM card (if any). Tables 3 and Table 4 present summary statistics from the perspective of teams and players, respectively.

The upper panel of Table 3 shows that approximately half of all auctions (156 out of 317) involved an incumbent. The probability that the player is sold is equal to 58% and 19% in auctions with and without incumbent, respectively. Among auctions with an incumbent, the probability that the incumbent becomes the winner is equal to 24% while the corresponding probability for a given non-incumbent is 5%. In 75 of these auctions the incumbent was eligible to use a RTM card and did use it in 17% of these cases, and in the remaining 81 auctions he did not because he did no longer possess a RTM card. As indicated in the middle panel, a team purchased on average 15 players through the auctions considered, comprising approximately of 3-4 batsmen, 6-7 bowlers, 1-2 wicket keepers, and 3-4 all-rounders. Furthermore, about 10 of the newly purchased players were Indian, and 7 newcomers. Teams bought on average 1.62 players through the RTM option. The lower panel shows that teams retained on average 3 players from their previous year's squad (these players do not appear in the auctions but have an indirect impact through the induced constraints in the auction), and were allocated an average budget of 5.65 million USD for purchasing players. On average, bidders consume 90% of their allocated budget. Last, there are no RTM cards that are left unused by the bidders at the end of the auction sequence.

Table 4 contains summary statistics on the auction data, for the full sample in the upper panel, and for the players from the performance sample in the lower one. The lower panel also reports statistics on our composite measure of performance and the wage earned by the cricket players. Conditional on observing bids at or above the reserve price, the average number of participants (i.e. the number of bidders having raised their paddle at some moment) is 2.3. The reserve price for sold players is found to be not significantly different compared to the sample average, in both cases it was set around 0.1 million USD. The fraction of newcomers and the within-set order of player appearance in the auctions is also similar across the two samples. The fraction of Indian players and the fraction of players eligible for RTM are, however, slightly higher in the sold-players sample. The average winning price is 0.33 million USD. There is actually a huge heterogeneity in the wages: the ratio between the highest and the lowest wage obtained in the auctions is as large as 140. Finally, we see that the performance score on average equals 23.9, and there is much dispersion in this variable as well since its standard deviation equals 14.55.

7.4 Empirical issues

In this section we introduce the empirical specifications chosen in our application. We also outline how our methodology is (slightly) adapted to take into account several features of the environment.

Again, we adopt a symmetric Gaussian structure and we also assume that $L = 1$ in A9. The variables we include in the vector of team/player characteristics $x_{i,f}$ are cricket-speciality dummies indicating whether i is of a certain speciality,[51] a dummy indicating whether i is of Indian nationality, a dummy indicating whether he is a newcomer, and set-specific fixed effects. We also include two indicator variables: *Bidder is incumbent* which corresponds to the variable $x_{i,f}^{inc}$ defined earlier and *Incumbent Present* which is equal to 1 if one of the eight firms is the incumbent and 0 otherwise (which thus corresponds to $\sum_f x_{i,f}^{inc}$). The vector of auction variables $z_{i,f}$ contains the order of sale of i within the set, the remaining budget of team f just before i is being auctioned, and five backlog variables: *# Batsman bought*, *# Bowlers bought*, *# Wicket-keepers bought*, *# All-rounders bought* and *# Overseas players bought*. Each of these variables is defined as the interaction between a variable counting the number of players of a given type already bought by f prior to the auction of i (including retained players), and a dummy indicating whether i is of this type. Finally, we include the variable *Incumbent present & no RTM card* which is equal to 1 if one of the eight firms is the incumbent and the incumbent has no RTM card, 0 otherwise. This variable aims to capture in a reduced form the dynamic effects coming from the RTM option: due to the scarcity of RTM cards, exerting the RTM option or not induces different continuation values throughout the auction sequence and thus modifies current bidding incentives.

To facilitate the maximisation of the likelihood function, we exploit that the drop-out price of the incumbent is always observed in our data: we take it as the last price at which the incumbent has raised its paddle. The presence of the RTM option for some of the IPL auctions requires also to adapt slightly our analysis. If we abstract from the fact that there is a limited number of RTM cards so that the use of a card is costly, then the optimal bidding strategy of the incumbent consists in remaining silent in the auction and then in using the RTM card if and only if her bid-signal $bs_{f^{inc}}$ is larger than $\widehat{bs}_{f^{inc}}(w)$ where w is the termination price. At a given price w, non-incumbents should decide whether or not to remain active as if they knew that the incumbent has a bid-signal below $\log(w) - A_{f^{inc}}$ (which reflects that the incumbent has not used her RTM card at w). Then, the corresponding cutoff bid-signal of a non-incumbent f, denoted by $\widehat{bs}_f^{RTM}(w)$, can be expressed by[52]

$$\widehat{bs}_f^{RTM}(w) = \log(w) - A_f - \frac{1}{(1-\theta)} \cdot \log\Big(\int_{-\infty}^{\log(w)-A_{f^{inc}}} e^{(1-\theta)\cdot U_{f^{inc}}(x)} d\frac{G_{f^{inc}}(x)}{G_{f^{inc}}(\log(w)-A_{f^{inc}})} \Big). \tag{21}$$

Consequently, accounting for RTM requires only small modifications in the likelihood function and the control functions. From the perspective of the identification of the bidding model,

[51] Included are dummies for batsman and bowler. There are too few wicket-keepers in the sold-player sample to add yet another speciality dummy.

[52] Under RTM, the characterization of a separating equilibrium as in Proposition 3.1 requires to strengthen a bit A3 to guarantee that the cutoff bid-signal is increasing in w. Intuitively, the strength of the informational externality should not be too large.

note that the RTM option would simplify a lot the argument: conditional on the observable covariates, the bid-signal of the incumbent is distributed independently of the auction termination price (which is determined from the competition between non-incumbents). Furthermore, since the incumbent should remain silent until the termination price is reached, then the decision of non-incumbents whether to remain active or not depends solely on their own bid-signal. Finally, given our independence assumption A1 ii), the drop-out times of the non-incumbents are independently distributed and we can then apply Athey and Haile (2002) to recover the bid distribution of the non-incumbents.

7.5 Estimation results

Table 5 contains our empirical results. The first column gives OLS estimates of the parameters appearing in the uncorrected performance equation (1). The numbers in brackets are the 95% CIs based on the usual OLS standard errors. The results indicate that the logarithm of the player's wage has a positive and statistically significant (at the 1% level) effect on his composite performance measure. A 1% increase in a player's wage is associated with a performance increase of 0.06 points. Alternatively a one standard deviation increase in the wage increases player's performance by 0.4 standard deviations. Regarding the team/player characteristics, we see that the indicators for Indian players, newcomers, and batsmen are not significant. The dummy indicating whether a player is a bowler is, however, significant and has a positive effect on performance. The two indicator variables 'Incumbent present' and 'Winner is incumbent' are not significant. As explained in Section 2, all these estimates are potentially biased because of sample selection and omitted variables. To correct for the bias, we now apply our correction method and report the results of the two estimation stages in columns 2 and 3.

Column 2 contains the first stage results, i.e., the ML estimates of the auction primitives $\beta^{inc,*}$, β^*_x, β^*_z, θ, σ_{PV}, and σ_{CV}, and the 95% CIs based on the asymptotic ML standard errors. In the discussion below, what is referred to as the *bid-valuation* of a team f for a player i corresponds to $\exp(A_{i,f} + bs^{PV}_{i,f} + bs^{id}_{i,f} + bs^{CV}_i + bs^{co}_i)$, i.e. the bidding amount that a firm is prepared to bid if she could observe the common value signals bs^{CV}_i and bs^{co}_i.

The lower panel of the table shows that the first stage estimate of θ is 0.77 (significant at the 1% level). Furthermore, the estimates of σ_{PV} and σ_{CV} are 1.48 and 2.22, indicating that a standard deviation increase in the aggregate bid-signal of a non-incumbent (resp. the incumbent) increase its bid-valuation by 0.44 (resp. 0.79) standard deviations.[53] More than half of the variance of bid-valuations comes thus from teams' private signals.

All team/player characteristics are statistically significant except the newcomer indicator:

[53]Bid-valuations are not observed and can neither be recovered for a specific player from our first-stage estimates. Nevertheless, what is identified from the first stage allow us to identify the standard deviation of the bid-valuations as $\sqrt{\sigma^2_{b-val} + \sigma^2_{PV} + \sigma^2_{CV}}$ with σ_{b-val} being the standard deviation of $\beta^*_f + \beta^*_x \cdot x_{i,f} + \beta^{inc,*} \cdot x^{inc}_{i,f} + \beta^*_z \cdot z_{i,f}$. This leads to the estimate 3.4.

teams reduce their bid-valuation for Indian players and increase it for certain player specialties (batsmen, bowlers); our results also indicate that bid-valuations are larger for players who belonged to one of the eight teams prior to the auctions (Incumbent present) and are even larger for the incumbent team (Bidder is incumbent).[54] Specifically, a player who participated in the tournament previously significantly increases the associated teams' bid-valuation by 0.75 standard deviations relative to other players when his incumbent team has still a RTM card.[55] This is consistent with the observation made earlier that the probability for a player to be sold is three times larger when there is an incumbent. Furthermore, the average difference between the incumbent's bid-valuation and the bid-valuation of a non-incumbent corresponds to 0.4 standard deviations, which is consistent with the observation that the incumbent's winning probability is five times larger than for a given non-incumbent.

Let us next look at the results concerning the auction variables z. The order of sale within the set does not significantly affect teams' bid-valuations. The coefficient associated with the logarithm of the remaining budget is highly significant, and is as expected positive, implying that teams bid more aggressively when they have more money to spend: a 10% increase in a team's remaining budget at a given point in the auction sequence, increases its bid-valuation for the given player by 0.015 standard deviations. The five last auction variables are our backlog variables capturing the bidder's past purchase behavior. Three of the backlog variables have, as one might have anticipated, negative and statistically significant impacts: an additional batsman (resp. bowler) acquired by a team reduces its bid-valuation for such a player by 0.10 (resp. 0.14) standard deviations; the reduction for an additional overseas player is 0.18 standard deviations, reflecting the constraint imposed by the auction organizers on non-Indian cricket players. The variable *# Wicket-keepers bought* is not significant, and *# All-rounders bought* is significant but somewhat surprisingly positive.

Column 3 reports the second stage results, i.e., the estimates of all parameters appearing in the augmented performance model (20), together with CIs based on standard errors obtained by a percentile bootstrap method using 1,000 bootstrapped samples. Using our control function approach, we find that the effect of wages is still significant (albeit now only at the 10% level), but smaller in magnitude: a one standard deviation increase in wages leads to an increase of performance by less than 0.3 standard deviations, that is to say only two thirds of the effect estimated by uncorrected OLS. This confirms that, as predicted in Section 2, and in accordance with our Monte Carlo results, naive OLS estimation leads to an upwards bias of the wage effect. The fact that wages still matter in explaining performance, even after controlling for sample selection and omitted variables, is (weak) evidence in support of theories emphasizing the

[54]Note that the coefficient $\beta^{inc,*}$ does not reflect the average difference between the logarithm of the bids of the incumbent and a non-incumbent: the incumbency status drives bids not only through the bid-valuations but also through the asymmetric information across bidders w.r.t. commonly valued signals and possibly through the RTM option.

[55]Given the estimated coefficient on the auction variable 'Incumbent present & no RTM card', the corresponding increase when the incumbent team has exhausted her RTM cards is only 0.37 standard deviations.

benefits of wages over market-clearing prices (Shapiro and Stiglitz (1984), Akerlof and Yellen (1990)).

The estimated effects of the team/player characteristics are of the same sign as the OLS estimates reported in column 1 and, as predicted in Section 2, are larger in magnitude except for the Indian player indicator. However, the implications of the significance tests do not change much either compared to those reported earlier. A notable exception is the variable indicating whether the player is matched with his incumbent team: its estimated effect has sharply declined relatively to the naive OLS estimate, and the variable is now statistically significant (at the 5% level). In line with what we predicted in Section 2, OLS thus indeed leads to an estimated impact of this variable which is biased towards zero. Note that, comparing columns 2 and 3, the incumbent indicator also happens to be the only variable for which the estimated coefficients of β_x^* and β_x are of different signs:[56] while players perform less well ceteris paribus when they are re-hired by their previous employer, the incumbent nonetheless values such players more highly. Using the expression of α^{inc} in terms of parameters that are identified from our first and second stages, our estimate for α^{inc} is 3.4 and is signicant at 5% level. It reflects that teams value their previous players for other reasons than their contribution to performance.

From the lower panel of the table we see that the coefficient on the control function CF^{CV} is statistically significant (at the 5% level with a one-sided test), but the one associated with CF^{PV} is not. The fact that (at least) one of our control functions is significant confirms that endogeneity is an issue in our application.

8 Conclusion

This paper develops a novel approach to consistently estimate the effects of contract features and principal-agent characteristics when the matching of principals to agents is determined by an auction. Our methodological approach consists mainly in introducing unobserved signals that jointly drive the auction and the post-auction outcomes, in developing an auction model where bidders anticipate the incentive effect of the auction price on the post-auction outcome, and last, in imposing restrictions guaranteeing that our interdependent value auction model is non-parametrically identified. We then propose to correct for the bias associated to the underlying multiple sources of endogeneity by using micro-founded control functions. Our methodology would also allow to develop rich counterfactual exercises: the joint estimation of the bidding and performance models allow to simulate the incidence on performance of different market rules or regulation, e.g. to analyze the benefits of excluding the incumbent.[57]

Beyond the labor and procurement applications, our approach can be extended and implemented in environments where selection emerges from competitive bidding. For instance, an

[56] *Newcomer* is the other exception but this variable is not significant in both cases.

[57] See Jehiel and Lamy (forthcoming) for a general theoretical result pleading for the exclusion of the incumbent.

important question in bankrupcy auctions is whether recovery rates are higher when the previous owner of the bankrupt firm (i.e., the incumbent according to our terminology) 'wins' the auction (Thorburn 2000). More generally, a large literature in empirical corporate finance has examined the post-acquisition outcomes of companies when they are acquired through an auction procedure (see Eckbo (2009) for a review).[58]

Another promising avenue for future research would be to analyse the relation between treasury or central bank[59] auctions and participating banks' post-auction performance, through a model that links structurally the auction and post-auction stages. Such auctions are ubiquitous in macro-finance and banking, but involve divisible or multiple goods instead of single ones. From an econometric perspective, this is challenging although some progress has been made especially for uniform and discriminatory auctions under pure private values (see Hortaçsu (2011) for a survey). In this vein, Cassola et al. (2013) show that the regression of a bank's performance/profitability on auction-based measures supposed to be proxies of a bank's short-term funding costs depends crucially on whether the auction-based measure is the final bid or the bank's willingness to pay (estimated by a structural approach).

References

ABOWD, J. M., F. KRAMARZ, AND D. N. MARGOLIS (1999): "High Wage Workers and High Wage Firms," Econometrica, 67, 251–333.

AKERLOF, G. A. AND J. L. YELLEN (1990): "The Fair Wage-Effort Hypothesis and Unemployment," Quarterly Journal of Economics, 105, 255–283.

ALLEN, J., R. CLARK, AND J.-F. HOUDE (2019): "Search Frictions and Market Power in Negotiated Price Markets ," Journal of Political Economy, 127, 1550–1598.

ARADILLAS-LOPEZ, A., A. GANDHI, AND D. QUINT (2013): "Identification and Inference in Ascending Auctions With Correlated Private Values," Econometrica, 81, 489–534.

ARYAL, G., I. PERRIGNE, AND Q. VUONG (2019): "Econometrics of Insurance Models with Multidimensional Types," mimeo.

ATHEY, S. AND P. HAILE (2002): "Identification of Standard Auction Models," Econometrica, 70, 2107–2140.

——— (2007): Nonparametric Approaches to Auctions, Amsterdam: North-Holland, chap. 60, 3847–3965, Handbook of Econometrics, Vol. 6A.

[58]Malmendier et al. (2012) consider competition between firms to merge with another firm. In general the sale procedure may not be a formal auction but rather a bargaining procedure that is modeled as an English auction.

[59]Central banks hold auctions for several purposes, including to purchase or repurchase assets (for quantitative easing), to sell or buy foreign exchange (for foreign exchange intervention) and sell collateral-backed short-term loans (for term auction facilities).

Athey, S. and J. Levin (2001): "Information and Competition in U.S. Forest Service Timber Auctions," Journal of Political Economy, 109, 375–417.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium," Econometrica, 63, 841–90.

Bhattacharya, V., A. Ordin, and J. W. Roberts (2018): "Bidding and Drilling Under Uncertainty: An Empirical Analysis of Contingent Payment Auctions," mimeo.

Bikhchandani, S., P. A. Haile, and J. G. Riley (2002): "Symmetric Separating Equilibria in English Auctions," Games and Economic Behavior, 38, 19–27.

Bikhchandani, S., S. Lippman, and R. Ryan (2005): "On the Right of First Refusal," The B.E. Journal of Theoretical Economics, 5, 1–44.

Bodoh-Creed, A. L. and B. R. Hickman (2018): "Pre-ollege Human Capital Investment and Affirmative Action: a Structural Policy Analysis of US College Admissions," Mimeo Haas School of Business and Queen's University.

Burchardi, K. B., S. Gulesci, B. Lerva, and M. Sulaiman (2019): "Moral Hazard: Experimental Evidence from Tenancy Contracts," The Quarterly Journal of Economics, 134, 281–347.

Cassola, N., A. Hortaçsu, and J. Kastl (2013): "The 2007 Subprime Market Crisis Through the Lens of European Central Bank Auctions for Short-Term Funds," Econometrica, 81, 1309–1345.

Chernozhukov, V. and H. Hong (2004): "Likelihood Estimation and Inference in a Class of Nonregular Econometric Models," Econometrica, 72, 1445–1480.

Chiappori, P.-A. and B. Salanié (2003): Testing Contract Theory: A Survey of Some Recent Work, Cambridge University Press, 115–149.

——— (2016): "The Econometrics of Matching Models," Journal of Economic Literature, 54, 832–61.

Choo, E. and A. Siow (2006): "Who Marries Whom and Why," Journal of Political Economy, 114, 175–201.

Cuesta, J. I. and A. Sepúlveda (2018): "Price Regulation in Credit Markets: A Trade-off between Consumer Protection and Credit Market," Mimeo University of Chicago.

Das, M., W. Newey, and F. Vella (2003): "Nonparametric Estimation of Sample Selection Models," Review of Economic Studies, 70, 33–58.

De Silva, D. G., G. Kosmopoulou, and C. Lamarche (2009): "The effect of information on the bidding and survival of entrants in procurement auctions," Journal of Public Economics, 93, 56–72.

DECAROLIS, F. (2014): "Awarding Price, Contract Performance, and Bids Screening: Evidence from Procurement Auctions," American Economic Journal: Applied Economics, 6, 108–32.

D'HAULTFOEUILLE, X. AND P. FÉVRIER (forthcoming): "The Provision of Wage Incentives : A Structural Estimation Using Contracts Variation," Quantitative Economics.

DUBIN, J. A. AND D. L. MCFADDEN (1984): "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption," Econometrica, 52, 345–62.

ECKBO, B. E. (2009): "Bidding strategies and takeover premiums: A review," Journal of Corporate Finance, 15, 149–178.

EINAV, L., A. FINKELSTEIN, S. P. RYAN, P. SCHRIMPF, AND M. R. CULLEN (2013): "Selection on Moral Hazard in Health Insurance," American Economic Review, 103, 178–219.

ENGELBRECHT-WIGGANS, R., P. R. MILGROM, AND R. J. WEBER (1983): "Competitive Bidding and Proprietary Information," Journal of Mathematical Economics, 11, 161–169.

GAYLE, G.-L. AND R. A. MILLER (2015): "Identifying and Testing Models of Managerial Compensation," Review of Economic Studies, 82, 1074–1118.

GIACOMINI, R., D. POLITIS, AND H. WHITE (2013): "A Warp-Speed Method for Conducting Monte Carlo Experiments Involving Bootstrap Estimators," Econometric Theory, 29, 567–589.

GNEEZY, U. AND J. A. LIST (2006): "Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments," Econometrica, 74, 1365–1384.

GOEREE, J. K. AND T. OFFERMAN (2003): "Competitive Bidding in Auctions with Private and Common Values," Economic Journal, 113, 598–613.

GOLLIER, C. (2001): The Economics of Risk and Time, MIT Press, San Diego.

GREENE, W. (2008): Econometric Analysis (sixth edition), Pearson.

GREENWALD, B. (1986): "Adverse Selection in the Labour Market," Review of Economic Studies, 53, 325–347.

GUERRE, E., I. PERRIGNE, AND Q. VUONG (2009): "Nonparametric Identification of Risk Aversion in First-Price Auctions Under Exclusion Restrictions," Econometrica, 77, 1193–1227.

HAILE, P. A., H. HONG, AND M. SHUM (2003): "Nonparametric Tests for Common Values at First-Price Sealed-Bid Auctions," NBER Working Paper No. w10105.

HECKMAN, J. J. (1979): "Sample Selection Bias as a Specification Error," Econometrica, 47, 153–61.

HENDRICKS, K., J. PINKSE, AND R. PORTER (2003): "Empirical Implications of Equilibrium Bidding in First-Price, Symmetric, Common Value Auctions," Review of Economic Studies, 70, 115–145.

HENDRICKS, K. AND R. H. PORTER (1988): "An Empirical Study of an Auction with Asymmetric Information," The American Economic Review, 78, 865–883.

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, 1415–1444.

HEUMANN, T. (2019): "An Ascending Auction with Multi-dimensional Signals," Journal of Economic Theory, 184.

HONG, H. AND M. SHUM (2003): "Econometric models of asymmetric ascending auctions," Journal of Econometrics, 112, 327–358.

HORTAÇSU, A. (2011): "Recent progress in the empirical analysis of multi-unit auctions," International Journal of Industrial Organization, 29, 345–349.

JACKSON, M. O. (2009): "Non-Existence of Equilibrium in Vickrey, Second-Price, and English Auctions," Review of Economic Design, 13, 137–145.

JEHIEL, P. AND L. LAMY (forthcoming): "On the benefits of set-asides," Journal of the European Economic Association.

JEHIEL, P. AND B. MOLDOVANU (1996): "Strategic Nonparticipation," Rand Journal of Economics, 27, 84–98.

JOFRE-BONET, M. AND M. PESENDORFER (2003): "Estimation of a Dynamic Auction Game," Econometrica, 71, 1443–1489.

KONG, Y. (2018): "Sequential Auctions with Synergy and Affiliation Across Auctions," mimeo.

KRASNOKUTSKAYA, E. (2011): "Identification and Estimation of Auction Models with Unobserved Heterogeneity," Review of Economic Studies, 78, 293–327.

KRISHNA, V. (2002): Auction Theory, Academic Press, San Diego.

LAFFONT, J.-J. AND J. TIROLE (1986): "Using Cost Observation to Regulate Firms," Journal of Political Economy, 94, 614–641.

——— (1987): "Auctioning Incentive Contracts," Journal of Political Economy, 95, 921–937.

——— (1993): A Theory of Incentives in Procurement and Regulation, The MIT Press.

LAFFONT, J.-J. AND Q. VUONG (1996): "Structural Analysis of Auction Data," American Economic Review (Papers and Proceedings), 86, 414–20.

LAMY, L., M. PATNAM, AND M. VISSER (2016): "Correcting for Sample Selection From Competitive Bidding, with an Application to Estimating the Effect of Wages on Performance," Tech. Rep. 11376, C.E.P.R. Discussion Papers.

LAZEAR, E. P. (2000): "Performance Pay and Productivity," American Economic Review, 90, 1346–1361.

LEE, D. AND N. G. RUPP (2007): "Retracting a Gift: How Does Employee Effort Respond to Wage Reductions?" Journal of Labor Economics, 25, 725–761.

LEWIS, G. AND P. BAJARI (2011): "Procurement Contracting With Time Incentives: Theory and Evidence," The Quarterly Journal of Economics, 126, 1173–1211.

LI, T., I. PERRIGNE, AND Q. VUONG (2000): "Conditionally independent private information in OCS wildcat auctions," Journal of Econometrics, 98, 129–161.

LU, J. AND I. PERRIGNE (2008): "Estimating risk aversion from ascending and sealed-bid auctions: the case of timber auction data," Journal of Applied Econometrics, 23, 871–896.

LUO, Y., I. PERRIGNE, AND Q. VUONG (2012): "Multiproduct Nonlinear Pricing: Mobile Voice Service and SMS," mimeo.

MAASLAND, E. AND S. ONDERSTAL (2007): "Auctions with Financial Externalities," Economic Theory, 32, 551–574.

MALMENDIER, U., E. MORETTI, AND F. PETERS (2012): "Winning by Losing: Evidence on Overbidding in Mergers," Review of Financial Studies, 31, 3212–3264.

MAS, A. (2006): "Pay, Reference Points, and Police Performance," Quarterly Journal of Economics, 121, 783–821.

MILGROM, P. AND R. WEBER (1982): "A theory of auctions and competitive bidding," Econometrica, 50, 1089–1122.

OYER, P. AND S. SCHAEFER (2011): Personnel Economics: Hiring and Incentives, Handbook of Labor Economics, Volume 4B. Great Britain: North Holland, 1769–1823.

PAARSCH, H. J. AND H. HONG (2006): An Introduction to the Structural Econometrics of Auction Data, MIT Press.

PAARSCH, H. J. AND B. SHEARER (2000): "Piece Rates, Fixed Wages, and Incentive Effects: Statistical Evidence from Payroll Records," International Economic Review, 41, 59–92.

PERRIGNE, I. AND Q. VUONG (2011): "Nonparametric Identification of a Contract Model With Adverse Selection and Moral Hazard," Econometrica, 79, 1499–1539.

PINKSTON, J. (2009): "A Model of Asymmetric Employer Learning with Testable Implications," Review of Economic Studies, 76, 367–394.

Powell, D. and D. Goldman (forthcoming): “Disentangling Moral Hazard and Adverse Selection in Private Health Insurance,” Journal of Econometrics.

Reny, P. J. (2011): “On the Existence of Monotone Pure-Strategy Equilibria in Bayesian Games,” Econometrica, 79, 499–553.

Reny, P. J. and S. Zamir (2004): “On the Existence of Pure Strategy Monotone Equilibria in Asymmetric First-Price Auctions,” Econometrica, 72, 1105–1125.

Roberts, J. W. (2013): “Unobserved heterogeneity and reserve prices in auctions,” The RAND Journal of Economics, 44, 712–732.

Shapiro, C. and J. E. Stiglitz (1984): “Equilibrium unemployment as a worker discipline device,” The American Economic Review, 74, 433–444.

Shearer, B. (2004): “Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment,” The Review of Economic Studies, 71, 513–534.

Thorburn, K. S. (2000): “Bankruptcy auctions: costs, debt recovery, and firm survival,” Journal of financial economics, 58, 337–368.

Vella, F. (1998): “Estimating Models with Sample Selection Bias: A Survey,” The Journal of Human Resources, 33, 127–169.

Weiergraeber, S. and C. Wolf (2018): “Bidder Asymmetries in Procurement Auctions: Efficiency vs. Information,” mimeo.

Wooldridge, J. (2015): “Control function methods in applied econometrics,” Journal of Human Resources, 50, 420–445.

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

Table 1: Monte Carlo results: Mean estimates of τ and β_x

	Mean of $\hat{\tau}$														
$\tau =$	-0.8			-0.5			0			0.5			0.8		
$\sigma_{CV} =$	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2
	N=300														
OLS	-0.22	0.16	-0.09	-0.02	0.33	0.20	0.31	0.59	0.63	0.64	0.82	0.88	0.85	0.95	0.97
PV	-0.78	-0.45	-0.41	-0.48	-0.21	-0.09	0.01	0.19	0.36	0.49	0.61	0.73	0.78	0.85	0.91
CV	-0.82	-0.81	-0.92	-0.53	-0.50	-0.62	-0.02	0.01	-0.11	0.48	0.49	0.39	0.79	0.78	0.71
	N=1,000														
OLS	-0.23	0.17	-0.10	-0.03	0.33	0.20	0.31	0.59	0.61	0.64	0.82	0.87	0.85	0.95	0.96
PV	-0.78	-0.47	-0.40	-0.49	-0.21	-0.09	0.01	0.20	0.36	0.50	0.61	0.72	0.79	0.85	0.89
CV	-0.82	-0.80	-0.93	-0.52	-0.51	-0.63	-0.02	0.00	-0.14	0.48	0.48	0.37	0.79	0.78	0.71
	Mean of $\hat{\beta}_x$														
$\tau =$	-0.8			-0.5			0			0.5			0.8		
$\sigma_{CV} =$	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2
	N=300														
OLS	0.42	0.10	0.34	0.52	0.23	0.35	0.68	0.46	0.43	0.84	0.72	0.66	0.92	0.88	0.83
PV	0.96	0.68	0.63	0.97	0.74	0.61	0.99	0.85	0.68	1.00	0.93	0.82	1.01	0.98	0.93
CV	1.05	1.08	1.14	1.04	1.06	1.12	1.03	1.02	1.09	1.01	1.01	1.06	0.99	1.01	1.02
	N=1,000														
OLS	0.43	0.09	0.35	0.52	0.22	0.35	0.68	0.46	0.44	0.84	0.73	0.66	0.93	0.89	0.84
PV	0.96	0.70	0.61	0.97	0.74	0.61	0.98	0.84	0.67	1.00	0.93	0.83	1.00	0.98	0.93
CV	1.04	1.07	1.13	1.04	1.06	1.12	1.03	1.03	1.11	1.01	1.02	1.06	1.01	1.01	1.02
Proportion Sold ($N = 1000$)	0.88	0.94	0.98	0.85	0.93	0.97	0.74	0.82	0.90	0.59	0.62	0.67	0.28	0.34	0.38

Table 2: Monte Carlo results: Power and bounds of confidence interval

								τ							
$\tau =$	-0.8			-0.5			0			0.5			0.8		
$\sigma_{CV} =$	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2
								N=300							
Power	1.00	0.94	1.00	0.96	0.74	0.93	0.05	0.05	0.10	0.91	0.80	0.70	0.86	0.69	0.95
Lower bound	-1.14	-1.25	-1.27	-0.82	-0.91	-1.03	-0.38	-0.37	-0.48	0.19	0.15	0.09	0.26	0.15	0.34
Upper bound	-0.54	-0.35	-0.60	-0.25	-0.13	-0.27	0.23	0.34	0.24	0.78	0.84	0.75	1.32	1.37	1.13
Empirical Coverage	0.94	0.96	0.86	0.94	0.94	0.89	0.95	0.95	0.9	0.96	0.94	0.92	0.96	0.95	0.94
								N=1,000							
Power	1.00	1.00	1.00	1.00	0.99	1.00	0.06	0.04	0.24	1.00	1.00	0.98	1.00	1.00	1.00
Lower bound	-0.99	-1.04	-1.14	-0.67	-0.72	-0.84	-0.17	-0.20	-0.33	0.32	0.30	0.20	0.54	0.50	0.51
Upper bound	-0.65	-0.57	-0.74	-0.37	-0.27	-0.45	0.11	0.19	0.07	0.64	0.69	0.56	1.04	1.08	0.91
Empirical Coverage	0.96	0.95	0.74	0.93	0.96	0.69	0.94	0.96	0.76	0.95	0.95	0.76	0.96	0.93	0.88
								β_x							
$\tau =$	-0.8			-0.5			0			0.5			0.8		
$\sigma_{CV} =$	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2
								N=300							
Power	1.00	0.98	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Lower bound	0.75	0.55	0.62	0.75	0.62	0.67	0.77	0.69	0.70	0.76	0.71	0.72	0.64	0.67	0.67
Upper bound	1.39	1.59	1.67	1.38	1.53	1.64	1.30	1.41	1.54	1.28	1.30	1.44	1.32	1.38	1.37
Empirical Coverage	0.92	0.95	0.91	0.94	0.94	0.91	0.94	0.94	0.92	0.95	0.95	0.96	0.96	0.94	0.96
								N=1,000							
Power	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Lower bound	0.87	0.82	0.85	0.89	0.81	0.87	0.89	0.84	0.89	0.88	0.85	0.88	0.83	0.82	0.83
Upper bound	1.22	1.34	1.44	1.20	1.31	1.41	1.17	1.22	1.32	1.15	1.18	1.24	1.18	1.20	1.20
Empirical Coverage	0.93	0.91	0.84	0.91	0.95	0.84	0.93	0.93	0.85	0.94	0.96	0.91	0.94	0.94	0.95

Table 3: Bidder summary statistics

Variable	Description		
		# Auctions	Percent of total
Bidder incumbency status:			
# Auctions without incumbent	Number of auctions without incumbent or with an incumbent who is not eligible to bid	161	50.79
# Auctions with incumbent & RTM	Number of auctions where player's previous team is eligible to use a RTM card	75	23.66
of which: # Auctions where RTM used	*Number of auctions where player's previous team uses RTM option*	*13*	*4.10*
# Auctions with incumbent & no RTM	Number of auctions where player's previous team is not eligible to use a RTM card	81	25.55
# Auctions with players sold	Number of auctions where player is sold to a team	122	38.49
# Auctions with incumbent and players sold	Number of auctions with an incumbent and where player is sold to a team	91	28.70
of which: # Auctions with players sold to incumbent team	*Number of auctions where player is sold to incumbent team*	*37*	*11.66*
		Mean	Std. Dev.
Bidder purchases in the auctions:			
# Players	Number of players bought in the auctions	15.25	2.49
speciality: # Batsman	Number of batsmen bought in the auctions	3.75	1.98
speciality: # Bowler	Number of bowlers bought in the auctions	6.5	2.07
speciality: # Wicket-Keeper	Number of wicket-keepers bought in the auctions	1.62	0.74
speciality: # All-Rounder	Number of all-rounders bought in the auctions	3.37	1.59
Nationality: # Indian	Number of Indian players bought in the auctions	10.12	1.72
# Newcomers	Number of players bought in the auctions who are newcomers	6.5	1.60
# RTM Bought Players	Number of players bought in the auctions through RTM	1.62	0.74
Bidder constraints in the auctions			
# Retained Players	Number of players retained by teams before the auctions	3	1.85
# RTM cards	Number of RTM cards received by teams	1.62	0.74
# of players where incumbent	Number of players eligible to be bought back using RTM (per team) if he has a RTM card	22.37	4.56
Spending cap	Amount of money allocated to a team	5.65	2.14
Remaining budget	Unused budget of a team at the end of the auctions	0.60	0.42

Note: All monetary values are reported in millions of USD. The currency used for the 2014 auctions was Indian Rupees (INR); we convert them to USD using an approximate conversion rate of 1 (USD) to 62 (INR).

Table 4: Summary statistics on auction and performance data

Variable	Description	Mean	Std. Dev.
Full sample:		N=317	
# active bidders	Participating bidders for each player auction	0.87	1.29
Reserve price	Reservation wage set by auctioneer	0.10	0.09
Order	Within-set order of player appearance in auction	5.23	2.80
Indian	Dummy indicating whether player is Indian	0.44	0.50
Newcomer	Dummy indicating whether player is a newcomer	0.32	0.47
Auction with incumbent	Dummy indicating whether one team is an incumbent (i.e. $\sum_f x_f^{dum}$)	0.49	0.50
of which RTM	Dummy indicating whether the incumbent is eligible to use a RTM card	0.23	0.42
Performance sample:		N=105	
# active bidders	Participating bidders for each player auction	2.27	1.07
Winning price	Equal to the final wage of the player	0.33	0.36
Reserve price	Reservation wage set by auctioneer	0.12	0.10
Order	Within-set order of player appearance in auction	5.28	2.90
Indian	Dummy indicating whether player is Indian	0.66	0.48
Newcomer	Dummy indicating whether player is a newcomer	0.42	0.49
Auction with incumbent	Dummy indicating whether one team is an incumbent (i.e. $\sum_f x_f^{dum}$)	0.74	0.43
of which RTM	Dummy indicating whether the incumbent is eligible to use a RTM card	0.32	0.47
Winner is incumbent	Dummy indicating whether player was matched with the winner in the previous season (formally it corresponds to $x_{f^{inc}}^{dum}$)	0.30	0.46
Tournament Points[†]	Player performance in tournament	23.90	14.55

Note: All monetary values are reported in millions of USD. The currency used for the 2014 auctions was Indian Rupees (INR); we convert them to USD using an approximate conversion rate of 1 (USD) to 62 (INR). † Construction of the composite performance measure is described in the Appendix. Note that the discrepancy between the size of the sold-player sample for performance (105) and the number of auctions where player is sold to a team (122) results from the fact that some players never perform on the field.

Table 5: Empirical results

	OLS	First Stage	Second Stage
Wage (τ)	5.80***		4.14**
	[2.50, 9.07]		[0.15, 8.11]
Team/player characteristics (x):			
Indian	-4.68	-2.52***	-3.22
	[-10.66, 1.30]	[-3.35, -1.68]	[-10.31, 4.48]
Newcomer	-1.17	0.51	-2.43
	[-9.54, 7.20]	[-0.10, 1.13]	[-11.58, 5.71]
Speciality: Batsman	2.16	2.16***	3.82
	[-5.76, 10.08]	[1.12, 3.20]	[-3.02, 11.88]
Speciality: Bowler	8.52**	3.01***	10.15***
	[1.64, 15.39]	[2.01, 4.02]	[3.47, 17.97]
Incumbent present	1.87	2.57***	5.48
	[-4.57, 8.31]	[1.27, 3.87]	[-3.05, 15.49]
Bidder is incumbent		2.23***	
		[1.23, 3.21]	
Winner is incumbent	-2.27		-15.27**
	[-8.26, 3.71]		[-30.80, -1.47]
Auction variables (z):			
Incumbent present & no RTM card		-1.33***	
		[-2.01, -0.65]	
Order of sale		-0.01	
		[-0.06, 0.05]	
Remaining budget (in logs)		0.55***	
		[0.29, 0.82]	
# Batsman bought		-0.36***	
		[-0.58, -1.14]	
# Bowlers bought		-0.48***	
		[-0.64, -0.32]	
# Wicket-keepers bought		0.01	
		[-0.41, 0.42]	
# All-rounders bought		0.41***	
		[0.24, 0.58]	
# Overseas players bought		-0.63***	
		[-0.80, -0.46]	
Other Structural Parameters:			
γ^{PV}			-0.47
			[-3.82, 2.05]
γ^{CV}			4.01**
			[0.30, 16.80]
θ		0.77***	
		[0.60, 0.94]	
σ_{PV}		1.48***	
		[0.75, 2.92]	
σ_{CV}		2.22***	
		[1.72, 2.87]	

Note: All specifications account for fixed effects with respect to the set in which the player was auctioned. Column 1 reports OLS estimates of the parameters in the performance equation (1), and 95% CIs based on the usual OLS standard errors. Column 2 gives the ML estimates of the auction model primitives β_x^*, $\beta^{inc,*}$, β_z^*, θ, σ_{PV}, and σ_{CV}, and 95% CIs based on the asymptotic ML standard errors. Column 3 reports OLS estimates of the parameters in the augmented performance equation (10), and 95% CIs based on a percentile bootstrapped procedure (with 1,000 bootstrapped samples). * indicates significance at 10%; ** at 5%; *** at 1%. The level of significance is based on two-sided tests except for the parameters γ^{PV} and γ^{CV} where it is one-sided given the restriction A9.

A Appendix

A.1 Proof of Proposition 3.1

When we have reached case A (resp. B with the incumbent's dropout price p^*), then we stay in case A (resp. B with p^*) forever and the strategy described in Proposition 3.1 is a dominant strategy. The case which is less straightforward is when we are in case C so that the auction can either switch to case B or to stop.

We show below that the bidding strategy $[\widehat{bs}_f^C]^{-1}(.)$ corresponds to firm f's best response in case C taking as given the incumbent bidding strategy $[\widehat{bs}_{f\,inc}]^{-1}(.)$ and that the firm will bid according to the strategy $[\widehat{bs}_f^B]^{-1}(.;p^*)$ in the continuation games under case B. Let $\Pi_f^C(bs_f, w)$ denote firm f's expected payoff if his bid-signal is bs_f, if he remains active until price w and if the incumbent has entered the auction and is still active while bidding optimally when the incumbent is no longer active. Let $\Pi_f^B(bs_f, w)$ denote the continuation payoff of firm f if the incumbent has dropped out at price w, if there is another non-incumbent which is active a price w and if he bids optimally afterwards (possibly by dropping out immediately after the incumbent's exit). Let $p_f(w)$ denote the belief of firm f about the probability that the auction stops if the incumbent drops out at price w. Thanks to our full support assumption on bid-signals, we have $p_f(w) > 0$ for any $w > 0$. As we will see, the exact value $p_f(.)$ (which depends on the other non-incumbents' strategies) does not play any role to characterize firm f optimal bidding strategy. We have that

$$\Pi_f^C(bs_f, b) = \int_{\widehat{bs}_{f\,inc}(W^r)}^{\widehat{bs}_{f\,inc}(b)} \Pi_f(bs_f, [\widehat{bs}_{f\,inc}]^{-1}(x)) \cdot \frac{dG_{f\,inc}(x)}{1 - G_{f\,inc}(\widehat{bs}_{f\,inc}(W^r))}$$

where

$$\Pi_f(bs_f, w) := \left[w^\theta \cdot e^{(1-\theta)\cdot[A_{f\,inc} + bs_f + U_{f\,inc}(\widehat{bs}_{f\,inc}(w))]} - w \right] \cdot p_f(w) + \Pi_f^B(bs_f, w) \cdot (1 - p_f(w)).$$

By definition of $\widehat{bs}_f^C(.)$, we have $[[\widehat{bs}_f^C]^{-1}(bs_f)]^\theta \cdot e^{(1-\theta)\cdot[A_{f\,inc} + bs_f + U_{f\,inc}(\widehat{bs}_{f\,inc}([\widehat{bs}_f^C]^{-1}(bs_f)))]} = [\widehat{bs}_f^C]^{-1}(bs_f)$. Furthermore, if the incumbent drops exactly at $[\widehat{bs}_f^C]^{-1}(bs_f)$, then firm f's best strategy is then to drop out immediately which yields a null payoff. We have thus $\Pi_f^B(bs_f, [\widehat{bs}_f^C]^{-1}(bs_f)) = 0$. On the whole we have $\Pi_f(bs_f, [\widehat{bs}_f^C]^{-1}(bs_f)) = 0$.

To conclude the proof we show that the optimal bid is $[\widehat{bs}_f^C]^{-1}(bs_f)$, we establish next that $\Pi_f(bs_f, w) >$ (resp. $<$) 0 if $w <$ (resp. $>$) $[\widehat{bs}_f^C]^{-1}(bs_f)$ for any possible strictly positive value for $p_f(w)$. Since $\widehat{bs}_f^C(.)$ is increasing, we get that $w^\theta \cdot e^{(1-\theta)\cdot[A_{f\,inc} + bs_f + U_{f\,inc}(\widehat{bs}_{f\,inc}(w))]} - w >$ (resp. $<$) 0 if $w <$ (resp. $>$) $[\widehat{bs}_f^C]^{-1}(bs_f)$.

For $w < [\widehat{bs}_f^C]^{-1}(bs_f)$, we have $\Pi_f^B(bs_f, w) \geq 0$: firm f can always guarantee himself a positive payoff by dropping out immediately at the price w. On the contrary, for any $w > [\widehat{bs}_f^C]^{-1}(bs_f)$, we have $\Pi_f^B(bs_f, w) \leq 0$: firm f would like to drop out immediately and would raise a negative payoff if he wins the auction at price w. **Q.E.D.**

Remark: Given the equilibrium bidding strategies and the absence of atoms in the signal distributions, it is clear that if a bidder dropouts at price p, then the probability that another bidder dropouts immediately is null (in particular if the incumbent dropouts at p^*, then bidder $f \neq f^{inc}$ will dropout out only if $bs_f = \widehat{bs}_f(p^*, \widehat{bs}_{f^{inc}}(p^*)))$, so that bidding ties would never occur with positive probability on the equilibrium path.

A.2 Proof of Proposition 4.1

From Athey and Haile (2002), we know that in the English (button) auction with no reserve price, with at least two bidders and with no covariates, then the independent pure private value model is non-parametrically identified from the joint observation of the identity of the winner and the winning price.

Take a given firm f^*. Applying Athey and Haile (2002) on any given point $\{(x_f, z_f)\}_{f=1,...,F}$ in the support of the observable covariates and for a given set of participants S such that $|S| \geq 2$ and $f^* \in S$ (with arises with positive probability given A7), the (centered) distributions G_f^{PV}, for each $f \in S$, and the associated value for $(A_f^*)_{f \in S}$ are thus identified. Then from A7, there exists a family of associated covariates $(x_{f^*}[k], z_{f^*}[k])_{k=1,...,K}$ so that the matrix $(\mathbf{1}, x_{f^*}^1[k], z_{f^*}[k])_{k=1,...,K}$ is of full rank and such that those realizations of covariates belong to the support of the observables among the sub-sample of auctions solely made up of non-incumbents. From the corresponding vector $\{A_{f^*}^*[k]\}_{k=1,...,K}$ that is identified, we recover the vector of coefficients $(\beta_{f^*}^*, \beta_x^*, \beta_z^*)$.

In words, the rest of our identification argument works roughly as follows: 1) the comparison of the bid-distribution of a bidder in an environment without incumbent with his bid distribution when being the incumbent allows to identify the distribution G^{CV}; 2) the comparison of the bid-distribution of a bidder in an environment without incumbent with his bid distribution when facing an incumbent allows to identify the parameter θ; finally we recover the parameter $\beta_x^{inc,*}$.

We now use the sub-sample of auctions with an incumbent. More precisely, we use the sub-sample of auctions i with a given incumbent f^{inc} and a given non-incumbent f^* and such that $S_i = \{f^{inc}, f^*\}$ (which has a positive measure given A8). We restrict ourselves to auctions with two bidders for the following reason: with at least three bidders, we could no longer apply Athey and Haile (2002) to recover the distribution of the drop-out price of each bidder, and this because drop-out prices would not be independently distributed because those bids are not fixed ex-ante as a function of the private signal of the corresponding bidder but would rather depend on the bidding history as characterized in Proposition 3.1. By contrast, with two bidders, we have the following two properties: 1) the drop-out price of each bidder does not vary across the bidding history since the auction stops immediately after the first drop-out; 2) the drop-out prices of the two bidders are distributed independently since their private signals are distributed independently.

Using the aforementioned auctions with two bidders, we can identify the distributions of the drop-out prices at which bidders f^{inc} and f^* would quit the auction. For any given realization of the covariates on the support of the observables, we then identify the CDF $G_{f^{inc}}$ and the scalar $A_{f^{inc}}$ from the distribution of the drop-out price of the incumbent. Since $G^{PV}_{f^{inc}}$ has been identified from the sub-sample without incumbents, a standard deconvolution argument now implies, thanks to A6, that G^{CV} is identified as well: Formally, if we let $\widehat{H}$ denote the Fourier transform of the CDF H, namely we have a bijection between H and $\widehat{H}$ characterized by the relations $\widehat{H}(\xi) = \int_{-\infty}^{+\infty} H(x)e^{-i\xi x}dx$ and $H(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\widehat{H}(\xi)e^{+i\xi x}dx$. The Fourier transform of the CDF of the sum of two independent variables is the product of the Fourier transform of the two underlying variables. A6 implies then that $\widehat{G}_{f^{inc}}(\xi) = \widehat{G}^{PV}_{f^{inc}}(\xi)\cdot\widehat{G}^{CV}(\xi)$. Finally G^{CV} is formally characterized as a function of the CDF $G_{f^{inc}}$ and $G^{PV}_{f^{inc}}$ by $G^{CV}(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\frac{\int_{-\infty}^{+\infty}G_{f^{inc}}(x')e^{-i\xi x'}dx'}{\int_{-\infty}^{+\infty}G^{PV}_{f^{inc}}(x'')e^{-i\xi x''}dx''}\cdot e^{+i\xi x}d\xi$.

The identification of $A_{f^{inc}}$ for a given set of covariates enables us to identify $A_f = A^*_f + A_{f^{inc}} - A^*_{f^{inc}}$ for any f and the scalar

$$\beta^{inc,*} + \frac{1}{1-\theta}\cdot\left[\lambda\beta_{f^{inc}} + \alpha_{f^{inc}} + \log[E[e^{\lambda\epsilon_{i,f^{inc}}}]]\right] = \beta^{inc,*} + \beta^*_{f^{inc}} - \frac{1}{1-\theta}\log(\int_{-\infty}^{+\infty} e^{(1-\theta)x}dG^{CV}(x)) \quad (22)$$

where $\beta^*_{f^{inc}}$ has been identified (separately) from the bidding data without incumbent and where G^{CV} has also been identified above. (22) implies that if θ is identified, then $\beta^{inc,*}$ will be identified as well.

Let us now show that the parameter θ is identified from to the distribution of the drop-out price of firm f^* in the sample where f^* faces incumbent f^{inc} and for a given set of covariates yielding some bid shifters, say A_{f^*} and $A_{f^{inc}}$, that has been identified as shown above. Let $G^{CV}(.|\widetilde{x}, f^{inc})$ denote the CDF of $bs^{co} + bs^{CV}$ conditional on $bs_{f^{inc}} = \widetilde{x}$. Under A6, the associated PDF, denoted by $g^{CV}(.|\widetilde{x}, f^{inc})$, satisfies $g^{CV}(x|\widetilde{x}, f^{inc}) = g^{CV}(x)\cdot g^{PV}_{f^{inc}}(\widetilde{x}-x)/g_{f^{inc}}(\widetilde{x})$, which is identified since all densities on the right hand side have already been identified. Using this additional notation, (4) can be written as $U_{f^{inc}}(\widetilde{x}) = \frac{1}{1-\theta}\cdot\log\left[\int_{-\infty}^{+\infty} e^{(1-\theta)\cdot x}\cdot dG^{CV}(x|\widetilde{x}, f^{inc})\right]$. We get from (5) that the drop-out price p of the non-incumbent f^* with the bid-signal bs_{f^*} is characterized by the equation $bs_{f^*} = \log(p) - A_{f^*} - U_{f^{inc}}(\log(p) - A_{f^{inc}})$.[60] Conditional on the scalar θ, the equilibrium drop-out price of the non-incumbent f^* as a function of the sum of its bid-signals bs_{f^*} can thus be expressed as $\Lambda(bs_{f^*}, \theta)$, where the function Λ has been identified from our previous identification steps. We then use the following result (see Proposition 2 in Gollier (2001)):

Lemma 1. *If $\widetilde{Y}$ is a stochastic variable with $Var(\widetilde{Y}) \neq 0$, then the function $z \to \frac{1}{z}\log(E[e^{z\cdot\widetilde{Y}}])$ is (strictly) increasing on $(0, +\infty)$.*

[60] Assumption A3 guarantees that this equation has a unique solution.

Given the expression of $U_{f^{inc}}$ and Lemma 1, we get that $\Lambda(bs_{i,f},\theta)$ is increasing in θ. Take a given quantile $q \in (0,1)$ of the bid-distribution of bidder f^*, denoted by p^q. Since the bid function is increasing, it corresponds then to the bid associated to the quantile q of the signal-distribution $G^{PV}_{f^*}$, denoted by bs^q. We have shown above that the drop-out price distribution of f^* has been identified and thus p^q, while bs^q has been identified from the bidding data without incumbents. At most one value of θ can then be consistent with $p^q = \Lambda(bs^q,\theta)$ and θ is thus identified. **Q.E.D.**

A.3 Details on the control functions

Case 3: From Proposition 3.1 we know that $w_i = W^r_i$ and $f^w_i \neq f^{inc}_i$ if and only if $bs_{i,f^{inc}_i} \leq \widehat{bs}_{i,f^{inc}_i}(W^r_i)$, $bs_{i,f^w_i} \geq \widehat{bs}^A_{i,f^w_i}(W^r_i)$ and $bs_{i,f} \leq \widehat{bs}^A_{i,f}(W^r_i)$ for each $f \in S_i \setminus \{f^{inc}_i, f^w_i\}$. Given A1, we have that conditional on $\mathscr{I}$, bs_{i,f^w_i} (resp. bs_{i,f^{inc}_i}) is distributed according to the distribution $G^{PV}_{f^w_i}$ (resp. $G_{f^{inc}_i}$) truncated below $\widehat{bs}^A_{i,f^w_i}(W^r_i)$ (resp. above $\widehat{bs}_{i,f^{inc}_i}(W^r_i)$). Using the same arguments as in cases 1 and 2, we obtain

$$CF^{PV}_i[l] = \int_{\widehat{bs}^A_{i,f^w_i}(W^r_i)}^{+\infty} x^l \cdot \frac{dG^{PV}_{f^w_i}(x)}{1 - G^{PV}_{f^w_i}(\widehat{bs}^A_{i,f^w_i}(W^r_i))}$$

and

$$\begin{aligned} CF^{CV}_i[l] &= \int_{-\infty}^{\widehat{bs}_{i,f^{inc}_i}(W^r_i)} E_{x\sim G^{CV}(.|\tilde{x},f^{inc}_i)}[x^l] \frac{dG_{f^{inc}_i}(\tilde{x})}{G_{f^{inc}_i}(\widehat{bs}_{i,f^{inc}_i}(w_i))} \\ &= \int_{-\infty}^{\widehat{bs}_{i,f^{inc}_i}(W^r_i)} \Big[\int_{-\infty}^{+\infty} x^l \cdot \frac{g^{CV}(x) g^{PV}_{f^{inc}_i}(\tilde{x}-x)}{G_{f^{inc}_i}(\widehat{bs}_{i,f^{inc}_i}(W^r_i))} dx \Big] \cdot d\tilde{x}. \end{aligned}$$

Under the symmetric Gaussian structure and $L = 1$, we get the inverse Mills ratios:

$$CF^{PV}_i = \sigma_{PV} \cdot \frac{\phi\Big(\frac{\widehat{bs}^A_{i,f^w_i}(W^r_i)}{\sigma_{PV}}\Big)}{1 - \Phi\Big(\frac{\widehat{bs}^A_{i,f^w_i}(W^r_i)}{\sigma_{PV}}\Big)} \quad \text{and} \quad CF^{CV}_i = -\frac{\sigma^2_{CV}}{\sqrt{\sigma^2_{PV} + \sigma^2_{CV}}} \cdot \frac{\phi\Big(\frac{\widehat{bs}_{i,f^{inc}_i}(W^r_i)}{\sqrt{\sigma^2_{PV}+\sigma^2_{CV}}}\Big)}{\Phi\Big(\frac{\widehat{bs}_{i,f^{inc}_i}(W^r_i)}{\sqrt{\sigma^2_{PV}+\sigma^2_{CV}}}\Big)}. \tag{23}$$

To get (16), we develop the expressions of $\widehat{bs}^A_{i,f^w_i}(W^r_i)$ and $\widehat{bs}_{i,f^{inc}_i}(W^r_i)$ in (23). In particular, we develop the third term appearing in $\widehat{bs}^A_{i,f^w_i}(W^r_i)$, namely

$$U^A_{i,f^{inc}_i} = \frac{1}{(1-\theta)} \cdot \log(E[e^{(1-\theta)\cdot(bs^{co}_i + bs^{CV}_i)} | bs_{i,f^{inc}_i} \leq \log(W^r_i) - A_{i,f^{inc}_i}]).$$

The expectation appearing in $U^A_{i,f_i^{inc}}$ is of the form $E[e^{y_1}|y_2 \leq a]$ with (y_1, y_2) following a bivariate centered normal distribution with the corresponding standard deviations $\sigma_1 = (1-\theta)\cdot\sigma_{CV}$ and $\sigma_2 = (1-\theta)\cdot\sqrt{\sigma_{PV}^2 + \sigma_{CV}^2}$ and the coefficient of correlation $\rho = \frac{\sigma_{CV}}{\sqrt{\sigma_{PV}^2+\sigma_{CV}^2}}$, and with $a = (1-\theta)\cdot\big[\log(W_i^r) - A_{i,f_i^{inc}}\big]$. Conditional on $y_2 = a$, y_1 follows then a normal distribution with mean $\rho\cdot\frac{\sigma_1}{\sigma_2}\cdot a$ and variance $(1-\rho^2)\cdot\sigma_1^2$. We obtain then that $E[e^{y_1}|y_2 = a] = \exp(\rho\cdot\frac{\sigma_1}{\sigma_2}\cdot a + (1-\rho^2)\cdot\frac{\sigma_1^2}{2})$ and then that $E[e^{y_1}|y_2 \leq a] = E[\exp(\rho\cdot\frac{\sigma_1}{\sigma_2}y_2)|y_2 \leq a]\cdot\exp((1-\rho^2)\cdot\frac{\sigma_1^2}{2})$. Then using the formula of the expectation of the exponential of a truncated normal distribution (see chapter 24 in Greene (2008)), we obtain finally that

$$E[e^{y_1}|y_2 \leq a] = \frac{1-\Phi(\rho\cdot\sigma_1 - \frac{a}{\sigma_2})}{\Phi(\frac{a}{\sigma_2})}\cdot e^{\frac{\sigma_1^2}{2}}.$$

Replacing σ_1, σ_2, ρ and a by the aforementioned values, we get

$$U^A_{i,f_i^{inc}} = \frac{1}{1-\theta}\log\Big(\frac{1-\Phi(\frac{(1-\theta)\cdot\sigma_{CV}^2 - [\log(W_i^r) - A_{i,f_i^{inc}}]}{\sqrt{\sigma_{PV}^2+\sigma_{CV}^2}})}{\Phi(\frac{[\log(W_i^r)-A_{i,f_i^{inc}}]}{\sqrt{\sigma_{PV}^2+\sigma_{CV}^2}})}\Big) + \frac{(1-\theta)}{2}\cdot\sigma_{CV}^2.$$

Case 4: For $l = 1,\dots,L$, let $CF_i^{PV}(f)[l] := E[(bs^{id}_{i,f_i^w} + bs^{PV}_{i,f_i^w})^l|\mathscr{I}, f_i^{sh} = f]$ and $CF_i^{CV}(f)[l] := E[(bs_i^{co} + bs_i^{CV})^l|\mathscr{I}, f_i^{sh} = f]$ denote the analogs of control functions defined in Section 4, except that here we condition on the identity of the second-highest, denoted by f_i^{sh}. Recall that the identity of this bidder is not assumed to be observed by the econometrician in general. To calculate the conditional control functions $CF_i^{PV}(f)[l]$ and $CF_i^{CV}(f)[l]$, we distinguish two sub-cases: 4a) the incumbent is the second-highest bidder; 4b) the incumbent is not the second-highest bidder.

Case 4a ($f_i^{sh} = f_i^{inc}$): From Proposition 3.1 we know that $f_i^{sh} = f_i^{inc}$ if and only if $bs_{i,f_i^{inc}} = \widehat{bs}_{i,f_i^{inc}}(w_i)$, $bs_{i,f_i^w} \geq \widehat{bs}^C_{i,f_i^w}(w_i)$ and $bs_{i,f} \leq \widehat{bs}^C_{i,f}(w_i)$ for each $f \in S_i \setminus \{f_i^{inc}, f_i^w\}$. Given A1 and using similar arguments as for the first three cases, we obtain

$$CF_i^{PV}(f_i^{inc})[l] = \int_{\widehat{bs}^C_{i,f_i^w}(w_i)}^{+\infty} x^l\cdot\frac{dG^{PV}_{f_i^w}(x)}{1-G^{PV}_{f_i^w}(\widehat{bs}^C_{i,f_i^w}(w_i))} \text{ and } CF_i^{CV}(f_i^{inc})[l] = \int_{-\infty}^{+\infty} x^l\cdot\frac{g^{CV}(x)g^{PV}_{f_i^{inc}}(\widehat{bs}_{i,f_i^{inc}}(w_i)-x)}{g_{f_i^{inc}}(\widehat{bs}_{i,f_i^{inc}}(w_i))}dx.$$

Under the symmetric Gaussian structure, we get tractable expressions for the cutoff bid-signals $\widehat{bs}^C_{i,f_i^w}(w_i)$ (which is defined in Section 3.3): $\widehat{bs}^C_{i,f_i^w}(w_i) = \log(w_i) - A_{i,f_i^w} - U_{f_i^{inc}}(\widehat{bs}_{f_i^{inc}}(w_i))$ with

$$U_{f_i^{inc}}(x) = \frac{\sigma_{CV}^2}{\sigma_{PV}^2+\sigma_{CV}^2}\cdot[x + \frac{1-\theta}{2}\sigma_{PV}^2], \tag{24}$$

a formula which comes from the same kinds of computations as for case 3 where we compute $U^A_{i,f_i^{inc}}$. Finally, for $l = 1$, we get the expression

$$CF_i^{PV}(f_i^{inc})[1] = \sigma_{PV} \cdot \frac{\phi\Big(\frac{\sigma_{PV}}{\sigma_{PV}^2+\sigma_{CV}^2}\log(w_i) - \frac{1}{\sigma_{PV}}\cdot A_{i,f_i^w} + \frac{\sigma_{CV}^2}{\sigma_{PV}\cdot(\sigma_{PV}^2+\sigma_{CV}^2)}\cdot A_{i,f_i^{inc}} - \frac{1-\theta}{2}\cdot\frac{\sigma_{PV}\cdot\sigma_{CV}^2}{\sigma_{PV}^2+\sigma_{CV}^2}\Big)}{1-\Phi\Big(\frac{\sigma_{PV}}{\sigma_{PV}^2+\sigma_{CV}^2}\log(w_i) - \frac{1}{\sigma_{PV}}\cdot A_{i,f_i^w} + \frac{\sigma_{CV}^2}{\sigma_{PV}\cdot(\sigma_{PV}^2+\sigma_{CV}^2)}\cdot A_{i,f_i^{inc}} - \frac{1-\theta}{2}\cdot\frac{\sigma_{PV}\cdot\sigma_{CV}^2}{\sigma_{PV}^2+\sigma_{CV}^2}\Big)},$$

and

$$CF_i^{CV}(f_i^{inc})[1] = \frac{\sigma_{CV}^2}{\sigma_{PV}^2+\sigma_{CV}^2}\cdot[\log(w_i) - A_{i,f_i^{inc}}].$$

Analogously as in case 3, the control function $CF_i^{PV}(f_i^{inc})[1]$ depends on the variables $x_{i,f}$ and $z_{i,f}$, for both $f = f_i^w$ and $f = f_i^{inc}$.

Case 4b ($f_i^{sh} \neq f_i^{inc}$): This case occurs if either the incumbent did not enter the auction at all, or the incumbent has participated but quit the auction before the second highest bidder. Let $\pi^{NP}_{i,f_i^{inc}}(w_i, f_i^w, f_i^{sh})$ denote the probability that the incumbent has not entered the auction for worker i conditional on the observable auction outcome (w_i, f_i^w, f_i^{sh}). Let $\pi^P_{i,f_i^{inc}}(p|w_i, f_i^w f_i^{sh})$ denote the CDF of the drop-out price of the incumbent conditional on the auction outcome (w_i, f_i^w, f_i^{sh}) and conditional on the incumbent having entered the auction. Note that the support of the distribution $\pi^P_{i,f_i^{inc}}(.|w_i, f^w, f_i^{sh})$ is $[W_i^r, w_i]$. Equipped with these notation, given A1 and by iterated expectations, we now have in case 4b:

$$CF_i^{PV}(f_i^{sh})[l] = \pi^{NP}_{i,f_i^{inc}}(w_i, f_i^w, f_i^{sh})\cdot\int_{\widehat{bs}^A_{i,f_i^w}(w_i)}^{\infty} x^l\cdot\frac{d[G^{PV}_{f_i^w}(x)]}{1-G^{PV}_{f_i^w}(\widehat{bs}^A_{i,f_i^w}(w_i))}$$

$$+(1-\pi^{NP}_{i,f_i^{inc}}(w_i, f_i^w, f_i^{sh}))\cdot\int_{W_i^r}^{w_i}\Big[\int_{\widehat{bs}^B_{i,f_i^w}(w_i,p)}^{\infty} x^l\cdot\frac{d[G^{PV}_{f_i^w}(x)]}{1-G^{PV}_{f_i^w}(\widehat{bs}^B_{i,f_i^w}(w_i,p))}\Big]\cdot d\pi^P_{i,f_i^{inc}}(p|w_i, f_i^w f_i^{sh})$$

and

$$CF_i^{CV}(f_i^{sh})[l] = \pi^{NP}_{i,f_i^{inc}}(w_i, f_i^w, f_i^{sh})\cdot CF[3] + (1-\pi^{NP}_{i,f_i^{inc}}(w_i, f_i^w, f_i^{sh}))\cdot\int_{W_i^r}^{w_i} CF[4b;p]\cdot d\pi^P_{i,f_i^{inc}}(p|w_i, f_i^w f_i^{sh})$$

where $CF[3]$ corresponds to the expression of the control function $CF_i^{CV}[l]$ in case 3, and $CF[4a;p]$ corresponds to the expression of the control function $CF_i^{CV}[l]$ in case 4a but where the final price w_i has been replaced by p (namely, we have $CF[4a;p] = \frac{\sigma_{CV}^2}{\sigma_{PV}^2+\sigma_{CV}^2}\cdot[\log(p) - A_{i,f_i^{inc}}]$).

Letting $p_i^{sh}(\mathscr{I}, f)$ denote the probability that the identity of the second-highest bidder is f conditional on $\mathscr{I}$, we can now, by iterated expectation, relate the unconditional and conditional control functions:

$$CF_i^{PV}[l] = \sum_{f \in S_i \backslash f_i^w} p_i^{sh}(\mathscr{I}, f) \cdot CF_i^{PV}(f)[l] \quad \text{and} \quad CF_i^{CV}[l] = \sum_{f \in S_i \backslash f_i^w} p_i^{sh}(\mathscr{I}, f) \cdot CF_i^{CV}(f)[l].$$

I) Computation of the probability $p_i^{sh}(\mathscr{I}, f)$ for any $f \in S_i \backslash f_i^w$. For $w > W_i^r$ and $f^w, f^{sh} \in S_i$, we let $\pi_i^*(w, f^w, f^{sh})$ denote the density (from an ex-ante perspective just before the auction takes place) associated to the event that the auction for worker i stops at wage $w \geq W_i^r$ and that the highest and second-highest bidders are respectively f^w and f^{sh}.
If $f_i^{inc} = f^{sh}$, then given our equilibrium characterization in Proposition 3.1 we have:

$$\pi_i^*(w, f^w, f_i^{inc}) = \Big[\prod_{f \in S_i \backslash \{f^w, f_i^{inc}\}} G_f^{PV}(\widehat{bs}_{i,f}^C(w)) \Big] \cdot g_{f_i^{inc}}(\widehat{bs}_{i,f_i^{inc}}(w)) \cdot (1 - G_{f^w}^{PV}(\widehat{bs}_{i,f^w}^C(w))).$$

If $f_i^{inc} \neq f^w, f^{sh}$, then either the incumbent did not enter the auction at all, or the incumbent quits the auction before the second- highest bidder. From Proposition 3.1, the first possibility occurs if and only if $bs_{i,f_i^{inc}} \leq \widehat{bs}_{i,f_i^{inc}}(W_i^r)$, $bs_{i,f^w} \geq \widehat{bs}_{i,f^w}^A(w)$ and $\max_{f \in S_i \backslash \{f_i^{inc}, f^w\}} \{[\widehat{bs}_{i,f_i^{inc}}^A]^{-1}(bs_{i,f})\} = w$; while the second possibility occurs if and only if there exists a price $p \in [W_i^r, w]$ at which the incumbent dropped out (before firm f^{sh}), and $bs_{i,f_i^{inc}} = \widehat{bs}_{i,f_i^{inc}}(p)$, $bs_{i,f^w} \geq \widehat{bs}_{i,f^w}^B(w_i, p)$, $[\widehat{bs}_{i,f^{sh}}^B]^{-1}(bs_{i,f^{sh}}; p) = w_i$ and $bs_{i,f} \leq \widehat{bs}_{i,f}^B(w_i, p)$ for any $f \in S_i \setminus \{f_i^{inc}, f^w, f^{sh}\}$. We have then:

$$\pi_i^*(w, f^w, f^{sh}) = G_{f_i^{inc}}(\widehat{bs}_{i,f_i^{inc}}(W_i^r)) \cdot \Big[\prod_{f \in S_i \{f^w, f^{sh}, f_i^{inc}\}} G_f^{PV}(\widehat{bs}_{i,f}^A(w)) \Big] \cdot g_{f^{sh}}^{PV}(\widehat{bs}_{i,f^{sh}}^A(w)) \cdot (1 - G_{f^w}^{PV}(\widehat{bs}_{i,f^w}^A(w)))$$

$$+ \int_{W_i^r}^{w} \Big[[\prod_{f \in S_i \backslash \{f^w, f^{sh}, f_i^{inc}\}} G_f^{PV}(\widehat{bs}_{i,f}^B(w, p'))] \cdot g_{f^{sh}}^{PV}(\widehat{bs}_{i,f^{sh}}^B(w, p')) \cdot (1 - G_{f^w}^{PV}(\widehat{bs}_{i,f^w}^B(w, p'))) \Big] \cdot d[G_{f_i^{inc}}(\widehat{bs}_{i,f_i^{inc}}(p'))].$$

From Bayesian updating (if $\sum_{f' \in S_i \backslash \{f_i^w\}} \pi_i^*(w, f_i^w, f') \neq 0$), we get finally that

$$\pi_i^{sh}(\mathscr{I}, f) = \frac{\pi_i^*(w_i, f_i^w, f)}{\sum_{f' \in S_i \backslash \{f_i^w\}} \pi_i^*(w_i, f_i^w, f')}.$$

II) Computation of $\pi_{i,f_i^{inc}}^{NP}(w, f^w, f^{sh})$ and $\pi_{i,f_i^{inc}}^{P}(.|w, f^w, f^{sh})$ for any $f^w, f^{sh} \in S_i \setminus \{f_i^w\}$ with $f^w \neq f^{sh}$. Let g_i^* denote the joint density of the vector $bs^* = \{bs_f\}_{f \in S_i}$ in auction for worker i. Given independence, $g_i^*(bs^*) = \prod_{f \in S_i \backslash \{f_i^{inc}\}} g_f^{PV}(bs_f) \cdot g_{f_i^{inc}}(bs_{f_i^{inc}})$. Let $\mathscr{S}_i(w, f^w, f^{sh})$ denote the set of vectors bs^* in $R^{|S_i|}$ that lead to the auction outcome w, f^w, f^{sh} in the auction for worker i. Let $\mathscr{S}_i(bs_f = x) = \{bs^* \in R^{|S_i|} | bs_f = x\}$ and $\mathscr{S}_i(bs_f \leq x) = \{bs^* \in R^{|S_i|} | bs_f \leq x\}$. Next we use the notation $\int_S g_i^*(bs^*) d\, bs^*$ (with S being a subset of $R^{|S_i|}$) to denote the density associated to the realization of the event S. Equipped with these notation, Bayesian updating leads to:

$$\pi^{NP}_{i,f_i^{inc}}(w,f^w,f^{sh}) = \frac{\int_{\mathscr{S}_i(w,f^w,f^{sh})\cap\mathscr{S}_i(bs_{f_i^{inc}}\leq \widehat{bs}_{i,f_i^{inc}}(W_i^r))} g_i^*(bs^*)d\,bs^*}{\int_{\mathscr{S}_i(w,f^w,f^{sh})} g_i^*(bs^*)d\,bs^*}$$

and

$$\pi^{P}_{i,f_i^{inc}}(p|w,f^w,f^{sh}) = \frac{\int_{\mathscr{S}_i(w,f^w,f^{sh})\cap\mathscr{S}_i(bs_{f_i^{inc}}=\widehat{bs}_{i,f_i^{inc}}(p))} g_i^*(bs^*)d\,bs^*}{(1-\pi^{NP}_{i,f_i^{inc}}(w,f^w,f^{sh}))\cdot\int_{\mathscr{S}_i(w,f^w,f^{sh})} g_i^*(bs^*)d\,bs^*}.$$

Note that the incumbent does not participate if and only if $bs_{f_i^{inc}} \leq \widehat{bs}_{i,f_i^{inc}}(W_i^r)$ and that the incumbent quits the auction at price p if and only if $bs_{f_i^{inc}} = \widehat{bs}_{i,f_i^{inc}}(p)$.
To complete the computation, note that we get from Proposition 3.1 that:

$$\int_{\substack{\mathscr{S}_i(w,f^w,f^{sh})\cap\\ \mathscr{S}_i(bs_{f_i^{inc}}\leq\widehat{bs}_{i,f_i^{inc}}(W_i^r))}} g_i^*(bs^*)d\,bs^* = \Big[\prod_{f\in S_i\setminus\{f_i^{inc},f^w,f^{sh}\}} G_f^{PV}(\widehat{bs}^A_{i,f}(w))\Big]\cdot g^{PV}_{f^{sh}}(\widehat{bs}^A_{i,f^{sh}}(w))\cdot[1-G^{PV}_{f^w}(\widehat{bs}^A_{i,f^w}(w))]\cdot G_{f_i^{inc}}(\widehat{bs}_{i,f_i^{inc}}(W_i^r)),$$

$$\int_{\substack{\mathscr{S}_i(w,f^w,f^{sh})\cap\\ \mathscr{S}_i(bs_{f_i^{inc}}=\widehat{bs}_{i,f_i^{inc}}(W_i^r))}} g_i^*(bs^*)d\,bs^* = \Big[\prod_{f\in S_i\setminus\{f_i^{inc},f^w,f^{sh}\}} G_f^{PV}(\widehat{bs}^B_{i,f}(w,p))\Big]\cdot g^{PV}_{f^{sh}}(\widehat{bs}^B_{i,f^{sh}}(w,p))\cdot[1-G^{PV}_{f^w}(\widehat{bs}^B_{i,f^w}(w,p))]\cdot \frac{dG_{f_i^{inc}}(\widehat{bs}_{i,f_i^{inc}}(p))}{dp},$$

and

$$\int_{\mathscr{S}_i(w,f^w,f^{sh})} g_i^*(bs^*)d\,bs^* = \Big[\prod_{f\in S_i\setminus\{f_i^{inc},f^w,f^{sh}\}} G_f^{PV}(\widehat{bs}^A_{i,f}(w))\Big]\cdot g^{PV}_{f^{sh}}(\widehat{bs}^A_{i,f^{sh}}(w))\cdot[1-G^{PV}_{f_i^w}(\widehat{bs}^A_{i,f^w}(w))]\cdot G_{f_i^{inc}}(\widehat{bs}_{i,f_i^{inc}}(W_i^r))$$
$$+\int_{W_i^r}^{w}\Big[\prod_{f\in S_i\setminus\{f_i^{inc},f^w,f^{sh}\}} G_f^{PV}(\widehat{bs}^B_{i,f}(w,p))\Big]\cdot g^{PV}_{f^{sh}}(\widehat{bs}^B_{i,f^{sh}}(w,p))\cdot[1-G^{PV}_{f_i^w}(\widehat{bs}^B_{i,f^w}(w,p))]\cdot d[G_{f_i^{inc}}(\widehat{bs}_{i,f_i^{inc}}(p))].$$

B Appendix (NOT FOR PUBLICATION)

B.1 Format of the tournament and player performance measure

In the IPL, a match is generally completed in 3 hours. The match involves one team *batting* (striking the ball) while the opposing team bowls (delivers the ball), followed by the opposing team batting. The objective of the batting team is to post the maximum amount of score in a certain period of time by striking the ball. The team that posts the highest score wins the match. A batsman is a player who specializes in hitting or 'striking' the cricket ball in order to score runs. A bowler is a player who specializes in delivering the ball to a batsman and whose primary aim is to dismiss the batsman or concede minimal runs. A wicket-keeper is a batsman who holds a special position in the field; his role is to stand behind the batsmen and guard the 'wicket' when a team is bowling, similar to the role of a catcher in baseball. All-rounders are players who are specialized in, both, batting and bowling. The general composition of a cricket team is three specialist batsmen, four all-rounders, three specialist bowlers and a wicket-keeper. The player specialities are an important feature of our auction model, because teams are implicitly constrained to select and bid in a way that optimizes their team composition (i.e., they are unlikely to buy only bowlers).

Our composite performance measure is derived from various, batting and bowling statistics observed for each player during the tournament[61]. The first step in that process was to award points for each basic statistic accumulated by each player across every match of the tournament. The mapping of game-specific player statistics to points is given in Table 6.

Table 6: Conversion of performance statistics into points

Performance Statistic	**Points**
# Runs - a batsman's score from striking the ball in the tournament	1 base points for each run
# 50s - Number of times a batsman ended the match with a score equal to or above 50 in the tournament	25 bonus points for each 50
# 100s - Number of times a batsman ended the match with a score equal to or above 100 in the tournament	50 bonus points for each 100
# Wickets - Number of batsman dismissed by a bowler in the tournament	25 base point for each wicket
4 Wickets - Number of times in the tournament when a bowler dismissed 4 batsman in one match	40 bonus points for 4-wicket haul
5 Wickets - Number of times in the tournament when a bowler dismissed 5 batsman in one match	50 bonus points for 5-wicket haul

Next, given the time constraints inherent in the format of the game and its emphasis on the rate of scoring, the player's total *number of points* was adjusted by a speciality specific factor. For batsmen, the factor measures the batsman's relative *strike-rate* in the tournament,[62] as the higher the strike rate, the more effective a batsman is at scoring quickly. For bowlers, the factor measures the bowler's relative *economy-rate* in the tournament,[63] as the higher the economy rate, the more effective a bowler is at limiting the opposition's total score. Finally, each player's adjusted points are divided by the number of games they played in the tournament as batsman or bowlers so that player's are judged on a per-game basis.

[61] See https://bit.ly/2CvCB44 for a description of an algorithm that constructs a similar performance measure to compare and rank players in the tournament. Note that the performance measure distinguishes only batsmen and bowlers; since wicket-keepers and all-rounders either bat (wicket-keepers) or bat and bowl (all-rounders), their performance is accounted through their batting and/or bowling scores.

[62] A batsman's strike-rate is defined as the average score of a batsman per 100 balls faced. Formally, this is equal to [100*(Batsman score/# Balls faced)]. The batsman's factor is his strike-rate divided by the average strike-rate of other batsmen in the tournament

[63] A bowler's economy-rate is defined as the average score conceded by a bowler per 6 balls. Formally, this is equal to [Bowler Score/(# Balls delivered/6)]. The bowler's factor is his economy-rate divided by the average economy-rate of other bowlers in the tournament

B.2 Likelihood function

Here we derive the likelihood function of the auction data when there is no RTM. The likelihood depends on the parameters β_f^*, $f = 1, \ldots, F$, β_x^*, $\beta^{inc,*}$, β_z^*, θ, and the parameters characterizing the (parametrized) distribution functions G^{CV}, G_f^{PV}, and G_f. Note that the cutoff bid-signal functions $\widehat{bs}_f$ and $\widehat{bs}_f^k$, $k = A, B, C$, depend on all these parameters. For notational simplicity we will not explicitly index the cutoff bid-signals, the distribution functions, and the other entities defined below, by the parameter vector. Furthermore, we will only give the likelihood function for observations where an incumbent is present among the potential bidders. The likelihood for the sub-sample without an incumbent corresponds to the likelihood for a standard independent private value model and its form can be found in Paarsch and Hong (2006).

Let $P_{i,f_i^{inc}}(w)$ denote the ex-ante probability that the incumbent prefers not to employ worker i at wage w given the value of the parameters. We thus have $P_{i,f_i^{inc}}(w) = G_{f_i^{inc}}(\widehat{bs}_{i,f_i^{inc}}(w))$, where $\widehat{bs}_{i,f_i^{inc}}(w)$ is defined in (3), and $p_{i,f_i^{inc}}(w) = g_{f_i^{inc}}(\widehat{bs}_{i,f_i^{inc}}(w))$ is the corresponding density evaluated at the cutoff bid-signal of the incumbent. Using the letters A and B associated with the various cases described in Section 3, we similarly define $P_{i,f}^A(w)$ (resp. $P_{i,f}^B(w,p)$ for $p \in [W_i^r, w]$) as the probability that firm f ($f \neq f_i^{inc}$) prefers not to employ i at wage w, conditional on observing that the incumbent has not entered the auction (resp. has entered and dropped out at price p). We have thus $P_{i,f}^A(w) = G_f^{PV}(\widehat{bs}_{i,f}^A(w))$ and $P_{i,f}^B(w,p) = G_f^{PV}(\widehat{bs}_{i,f}^B(w,p))$, where the cutoff bid-signals $\widehat{bs}_{i,f}^A(w)$ and $\widehat{bs}_{i,f}^B(w,p)$ are defined in Section 3. We also define $P_{i,f}^C(w)$ as the probability that firm f does not wish to employ i at w, given that the incumbent is still active at w and that this firm believes the incumbent is going to quit instantly at this price. This probability can hence be written as $P_{i,f}^C(w) = G_f^{PV}(\widehat{bs}_{i,f}^C(w))$. Finally, let $p_{i,f}^k(w) := \frac{dP_{i,f}^k(w)}{dw}$ for $k = A, C$ and $p_{i,f}^B(w,p) := \frac{dP_{i,f}^k(w,p)}{dw}$, the associated densities.

For a given value of the parameter vector, the likelihood associated with the event that worker i remains unsold conditional on $\mathscr{I}$, is denoted L_i^{unsold}. Given Proposition 3.1 and the independence assumption A1, we have:

$$L_i^{unsold} = \prod_{\substack{f=1 \\ f \neq f_i^{inc}}}^{F} P_{i,f}^A(W_i^r) \times P_{i,f_i^{inc}}(W_i^r).$$

To write down the other terms of the likelihood function, we now use that the identity of the second-highest bidder is assumed to be observed by the econometrician. Letting f_i^{sh} be the identity of the the second-highest bidder in auction i (if there is not one, then $f_i^{sh} := 0$), it is thus assumed that $\mathscr{I}$ contains f_i^{sh} for all i. The likelihood associated with the event that i is sold to firm f_i^w at w_i, and the second highest bidder is f_i^{sh}, conditional on $\mathscr{I}$, is denoted $L_i^{sold}(w_i, f_i^w, f_i^{sh})$. The precise form of this type of likelihood contribution depends on whether the incumbent is the winner, the second highest bidder, or neither of these two bidders. It also depends on whether i is sold at or strictly above the reserve price.

If there is a single entrant (so that $w_i = W_i^r$ and $f_i^{sh} = 0$), we have:

$$L_i^{sold}(w_i, f_i^w, f_i^{sh}) = \prod_{\substack{f=1 \\ f \neq f_i^w}}^{F} P_{i,f}^C(W_i^r) \times \left(1 - P_{i,f_i^w}(W_i^r)\right) \text{ if } f_i^w = f_i^{inc},$$

and

$$L_i^{sold}(w_i, f_i^w, f_i^{sh}) = P_{i,f_i^{inc}}(W_i^r) \times \prod_{\substack{f=1 \\ f \neq f_i^{inc}, f_i^w}}^{F} P_{i,f}^A(W_i^r) \times \left(1 - P_{i,f_i^w}^A(W_i^r)\right) \text{ if } f_i^w \neq f_i^{inc}.$$

If there are at least two entrants (so that $w_i > W_i^r$ and $f_i^{sh} \neq 0$), we have:

$$L_i^{sold}(w_i, f_i^w, f_i^{sh}) = \prod_{\substack{f=1 \\ f \neq f_i^w, f_i^{sh}}}^{F} P_{i,f}^C(w_i) \times p_{i,f_i^{sh}}^C(w_i) \times \left(1 - P_{i,f_i^{inc}}(w_i)\right) \text{ if } f_i^w = f_i^{inc},$$

$$L_i^{sold}(w_i, f_i^w, f_i^{sh}) = \prod_{\substack{f=1 \\ f \neq f_i^w, f_i^{sh}}}^{F} P_{i,f}^C(w_i) \times p_{i,f_i^{sh}}(w_i) \times \left(1 - P_{i,f_i^w}^C(w_i)\right) \text{ if } f_i^{sh} = f_i^{inc},$$

and

$$L_i^{sold}(w_i, f_i^w, f_i^{sh}) = \prod_{\substack{f=1 \\ f \neq f_i^w, f_i^{sh}, f_i^{inc}}}^{F} P_{i,f}^A(w_i) \times p_{i,f_i^{sh}}^A(w_i) \times \left(1 - P_{i,f_i^w}^A(w_i)\right) \times P_{i,f_i^{inc}}(W_i^r)$$

$$+ \int_{W_i^r}^{w_i} \prod_{\substack{f=1 \\ f \neq f_i^w, f_i^{sh}, f_i^{inc}}}^{F} P_{i,f}^B(w_i, p) \times p_{i,f_i^{sh}}^B(w_i, p) \times \left(1 - P_{i,f_i^w}^B(w_i, p)\right) \times p_{i,f_i^{inc}}(p) dp \text{ if } f_i^{inc} \neq f_i^w, f_i^{sh}.$$

B.3 Warp-Speed Monte Carlo

The Warp-Speed method will be described by considering the parameter τ. The methodology is strictly the same for any other parameter. From the k-th replication sample ($k = 1, \ldots, 1,000$), we draw a single bootstrap sample of size N. Letting $\hat{\tau}_{N,k}$ and $\hat{\tau}_{N,k}^*$ be the second-stage estimates using the k-th Monte Carlo sample and its associated bootstrap resample, respectively, we can then construct a sequence of 95% confidence intervals for τ

$$CI_{N,k}(\tau) = [\hat{\tau}_{N,k} - q_N(0.975), \hat{\tau}_{N,k} - q_N(0.025)], \quad \text{and} \quad k = 1, \ldots, 1,000,$$

where $q_N(0.025)$ and $q_N(0.975)$ are the 0.025-quantile and 0.975-quantile of the empirical distribution of $\hat{\tau}^*_{N,k} - \hat{\tau}_{N,k}$, $k = 1, \ldots, 1,000$, respectively. We can now estimate the lower bound (resp. upper bound) of the 95% confidence interval of τ by taking the mean over the lower bounds (resp. upper bounds) of $CI_{N,k}(\tau)$, $k = 1, \ldots, 1,000$. Similarly, the power of the t-test of the null hypothesis $H_0 : \tau = 0$ (against the bilateral alternative) can simply be estimated by the fraction of times zero does not belong to $CI_{N,k}(\tau)$, $k = 1, \ldots, 1,000$, given that the data are generated under a particular value of τ. The novelty of the method proposed by Giacomini et al. (2013) is that only one bootstrap resample is required for each replication (instead of some high number as in a standard Monte Carlo experiment), thereby drastically reducing the computation time.