



**HAL**  
open science

## Distinguishing incentive from selection effects in auction-determined contracts

Laurent Lamy, Manasa Patnam, Michael Visser

► **To cite this version:**

Laurent Lamy, Manasa Patnam, Michael Visser. Distinguishing incentive from selection effects in auction-determined contracts. *Journal of Econometrics*, 2023, 235 (2), pp.1172-1202. 10.1016/j.jeconom.2022.10.003 . hal-04382099

**HAL Id: hal-04382099**

**<https://hal.science/hal-04382099>**

Submitted on 10 Jan 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Distinguishing Incentive from Selection Effects in Auction-Determined Contracts

Laurent Lamy      Manasa Patnam<sup>†</sup>      Michael Visser<sup>‡</sup>

January 10, 2024

## Abstract

This paper develops a novel approach to estimate how contract and principal-agent characteristics influence a post-auction outcome when the matching between agents and principals derives from an auction process. We propose a control-function approach to account jointly 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 the endogeneity biases well, even in small samples. We apply our methodology to a labor market application and estimate the effect of sports players' auction-determined wages on their individual performances. We also use our structural estimates to evaluate the strength of matching inefficiencies and assess counterfactual reservation wage policies.

Keywords: Econometrics of Auctions; Econometrics of Contracts; Endogenous Matching; Polychotomous Sample Selection; Price-performance Elasticity.

JEL classification: C34; C57; D44; M52; Z22.

## 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. 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 \(1984\)](#)). At the same time the wage increase also changes the pool of workers that are matched to the firm, in particular on the basis of unobserved

---

CIREA (Ecole des Ponts, ParisTech). Email: laurent.lamy78@gmail.com. Corresponding author.

<sup>†</sup>ENSAE-CREST. Email: mpatnam1@gmail.com.

<sup>‡</sup>ENSAE-CREST and CRED, University of Paris 2. Email: michael.visser@ensae.fr.

agent characteristics, including those that create idiosyncratic synergies (Abowd et al. (1999)). Similarly, an insurance company lowering the coverage of its auto-insurance contract could motivate agents to drive more prudently, but this may also push out of its portfolio of customers the riskiest agents. Both the incentive and selection effects of a less comprehensive coverage are thus expected to reduce the number of claims per insured agent that the insurer receives (Chiappori and Salanié (2003)).

The empirical literature mainly focuses on the “incentive effect” of contracts. Incentive 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 agents reciprocate (Gneezy and List (2006), Mas (2006) and Lee and Rupp (2007))<sup>1</sup>, or to measure to what extent explicit incentives schemes like piece-rates pay or sharecropping arrangements enhance output or productivity (Lazear (2000), Shearer (2004), Burchardi et al. (2019)).<sup>2</sup> Although quantifying incentive effects is admittedly interesting, a precise assessment of the matching process that drives how agents self-select into contracts is also important in practice, for instance to evaluate alternative counterfactual policies.

Disentangling incentive from selection effects is, however, a challenging issue as stressed by Chiappori and Salanié (2003). Most papers have addressed this issue 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 either of the contracts posted by the principal (d’Haultfoeuille and Février (2020), Powell and Goldman (2021)) or of the agent’s preferences/costs (Abito (2020)).<sup>3</sup> 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 observable production costs and the associated transfers).

In many environments, however, there are multiple principals who compete to be matched with various agents. This paper develops a methodology to consistently estimate how contract and principal-agent characteristics influence agents’ ex post outcome when the matching of agents to principals and a contract characteristic are determined through auctions. We present our methodology through the labor auction framework that we use later as an empirical illustration: to hire workers (or agents according to our terminology above), firms (the principals)

---

<sup>1</sup>These empirical studies consider very different kinds of work (e.g., data entry for a university library in Gneezy and List (2006), police investigation in Mas (2006), airline pilots in Lee and Rupp (2007)) and find that such behavioral effects tend to decline over time.

<sup>2</sup>Lazear (2000) argues that the productivity rise of a firm switching to performance-based contracts comes also from the fact that such contracts attract the most able workers.

<sup>3</sup>Abito (2020) exploits the institutional feature that regulation contracts for electric utilities remain fixed during a predetermined period while their costs and efforts change.

participate in various auctions which assign each worker to the winning bidder (if any) and where the winning bid determines the wage. For each worker matched to a firm, a performance measure is observed after the auction. This performance measure is allowed to depend directly on the auction-determined wage: This is a reduced-form approach to reflect what we refer to as the *incentive effect* and which can result from various channels. In our application, under-performing workers risk a termination of their contract after one year (see section 6.2 for institutional details), implying that better paid workers have an incentive to exert more effort, very much like in our opening example. Another channel is a behavioral reaction like reciprocity in response to a wage that is perceived as fair or unfair (Akerlof and Yellen (1990)). Our approach remains agnostic on the micro-foundation of the incentive effect.

The main features of our model are the following. Besides the auction-determined wage and various firm-worker observable characteristics, the post-auction performance is assumed to depend on two signals that are referred to as *productivity signals*. The first signal reflects match-specific synergies, observed privately by the firm concerned. The second signal reflects worker-specific attributes that are commonly valued by all firms, but observed only by the incumbent, i.e., the firm previously matched with the worker (if there is one). At the auction stage, firms are assumed to value a worker through their expectations regarding the post-auction performance of this worker and also through two additional signals, which we refer to as *intrinsic signals*. Adopting the same structure as for the productivity signals, one intrinsic signal is match-specific, i.e., idiosyncratic to each firm-worker pair, and one is worker-specific, i.e., is commonly valued by all firms but observed only by the incumbent. Regarding the auction rules, we consider that workers are auctioned through separate English auctions. A publicly observable reservation wage is also attached to each worker: it corresponds to the starting price in the auction. The equilibrium bidding functions of our auction model determine then the auction outcome, i.e., whether the worker is matched with a firm, and if so the wage and the identity of the corresponding “winning” firm, as a function of this full set of productivity and intrinsic signals.

The two productivity signals that enter the performance equation represent, from the econometrician’s viewpoint, an error term capturing unobserved worker/firm heterogeneity. As detailed in Section 2, they are a source of three kinds of endogeneity. First of all, they determine, together with the intrinsic signals, for which workers there is bidding above the reservation wages, and thereby the sample on which estimation is based. Second, if there is bidding above the reservation wage, the winning bid is partly determined by these two productivity signals, creating a correlation between the wage and the error term in the performance equation. Third, each worker is not assigned randomly to a firm, but to the specific firm that wins the auction. The process which determines the winning bidder’s identity is yet again partly determined by the two productivity signals entering the performance equation, implying that there is endogeneity in the matching between workers and firms. Standard estimation of our performance equation would therefore lead to biased parameter estimates.

To tackle the resulting endogeneity biases in the (post-auction) performance equation, we proceed in the following way. First, we characterize the perfect Bayesian equilibrium of the

bidding model. We show in particular that firms' bidding strategies are increasing functions of their aggregate signal, that is to say the sum of the signals they observe. Second, we provide conditions under which the distribution of the aggregate signals and the incentive effect are non-parametrically identified. Third, the identification of our auction model allows us to identify two types of control function terms for the performance equation: one type, referred to as the match-specific control terms, is associated with the match-specific signals of the winning bidder, while the other, referred to as the worker-specific terms, is associated with the worker-specific signals received by the incumbent. A Monte Carlo study shows that estimation of the performance equation by OLS leads to strongly biased estimates. On the contrary, our endogeneity-corrected estimator is shown to be unbiased even in relatively small samples.

We illustrate our methodology using data on cricket players participating in a tournament with a unique setting where team assignment and wage determination result from a sequence of English auctions. We estimate the incentive effect and also the *incumbency effect*, i.e., whether players perform better when staying in the same team *ceteris paribus*. We find important differences between OLS and our bias-corrected estimates: both estimates of the incentive effect are statistically significant but OLS leads to an upward bias; the incumbency effect is negative using both methods but is only statistically significant with our control function method. These discrepancies are also reflected in a large and significant value of the worker-specific control function term, confirming that endogeneity plays an important role in our data. Using our parameter estimates, we also evaluate the degree of match inefficiencies induced by the worker-specific attributes that are observed only by the incumbent and conduct counterfactuals to evaluate alternative reservation wage policies. This illustrates how our methodology can be used beyond the estimation of the (post-auction) performance equation.

While we frame our presentation through a labor auction model, our analysis is relevant beyond labor environments. In the concluding section, we give a more detailed description of how our method could be applied, in particular in auctions for procurement where competing firms are bidding on the contract characteristics that can induce moral hazard effects and thus impact post-procurement outcomes.<sup>4</sup> Furthermore, our methodology can be applied even if the matching mechanism is not explicitly an auction. In many matching markets (e.g. labor or housing), formal auctions are not frequently used, but if we are prone to abstract from search costs, auction models can be used as a first approximation to model informal search behavior in these markets.

Our work contributes to the structural empirical literature that considers the joint modeling of auction and post auction data (see [Perrigne and Vuong \(2021\)](#) for a review on the structural econometrics of auction data). In the seminal contributions of [Athey and Levin \(2001\)](#) and [Hendricks et al. \(2003\)](#), the linkage between the auction and the post auction data comes from the fact that bidders are privately informed about the composition of timber species and

---

<sup>4</sup>Disentangling moral hazard from selection effects received attention in the procurement literature: [Ryan \(2020\)](#) considers power procurement auctions in India and analyzes ex post renegotiation which is driven both by auction-determined contract characteristics and by the producer's idiosyncratic renegotiation abilities. [Lewis and Bajari \(2011\)](#) consider highway procurement in US and analyze actual project-delivery time which is driven by a component of the bid and the winning firm's idiosyncratic ability to complete the project on time.

the quantity of oil, respectively: informational asymmetries call for bidding cautiously to avoid the so-called winner's curse. More recent contributions develop models where the post auction realization is correlated not only to information held by the bidders but also to decisions made ex post by the auction winner (i.e., a kind of moral hazard) and that are driven not only by its privately observed type but also by its bid which determines part of the contract characteristics.<sup>5</sup> [Bhattacharya et al. \(2022\)](#) estimate the optimal royalty in oil lease auctions through a common value model where the unique source of inefficiency comes from ex post moral hazard (the drilling decision and its timing). [Kong et al. \(2022\)](#) also consider auctions for the option to drill an oil field but through a bi-dimensional private value model which opens the door to matching inefficiencies.<sup>6</sup> As in [Kong et al. \(2022\)](#), we consider a model where each bidder receives a private value signal that enters its valuation through the future performance (the match-specific productivity signal) and one that is not related to future performance (the match-specific intrinsic signal). On top of this structure, the worker-specific signals involve an additional form of asymmetry. Overall, our analysis departs from the literature on the structural econometrics of auction data for two main reasons: 1) our auction model involves interdependent values with asymmetric bidders; 2) we depart from the quasi-linear payoff paradigm where the auction price enters linearly bidders' payoff function.

We also contribute to the literature that uses a control function approach to correct for sample selection or endogeneous regressors. [Heckman \(1979\)](#)'s seminal contribution, extended to polychotomous outcomes by [Dubin and McFadden \(1984\)](#), deals mainly with selection.<sup>7</sup> The literature on production function estimation, where inputs depend on the firm's (unobserved) productivity shock, offers also a similar approach: [Olley and Pakes \(1996\)](#) and [Akerberg et al. \(2015\)](#) address the resulting endogeneity problem with a dynamic setup that captures capital accumulation and time-correlation of the productivity shocks but rules out unobserved heterogeneity across firms. However, the vast majority of papers in this literature adopt a reduced form approach to construct control functions (see [Vella \(1998\)](#) and [Wooldridge \(2015\)](#) for surveys). The control functions we derive are instead micro-founded by explicitly relying on auction theory and tackle the multiple sources of bias that occur simultaneously.<sup>8</sup>

---

<sup>5</sup>In a related vein, [Bodoh-Creed and Hickman \(2018\)](#) model the student-college match through an all-pay contest (with private values) where selection is based on the student's public academic score and an idiosyncratic matching shock. Their structural approach allows to disentangle the effects of college quality, unobserved student characteristics and pre-college human capital investment on earnings. An important technical difference with the auction literature is that [Bodoh-Creed and Hickman \(2018\)](#) consider the equilibrium of the limit game with a continuum of agents.

<sup>6</sup>Note that models involving multidimensional types receive also some attention in the empirical literature on contracts with a monopolistic principal we mentioned earlier. [Luo et al. \(2012\)](#) consider e.g. consumers for telecommunication services where the multidimensional types capture preferences for different products. In [Aryal et al. \(2019\)](#)'s model with bi-dimensional types, the agents are drivers who differ in terms of expected number of accidents and absolute risk aversion coefficient.

<sup>7</sup>Contrary to [Heckman \(1979\)](#) and [Dubin and McFadden \(1984\)](#) where matching is driven by a latent unobserved variable, [Akkus et al. \(2021\)](#) consider an assortative matching model where the value of each realized match is observed by the econometrician. The residual of the match value equation (which is estimated through a censored regression given that the value of unrealized matches are unobserved) is then used as a control in the equation of interest.

<sup>8</sup>[Wooldridge \(2002\)](#) and [Das et al. \(2003\)](#) propose a control function method to address jointly the problems of endogeneous regressors and sample selection with a parametric and non parametric approach, respectively.

The remainder of the paper is organized as follows. Section 2 presents the environment and sources of bias. Section 3 derives how firms and agents are matched through an auction model with interdependent values. Section 4 is devoted to the econometric methodology while Section 5 presents the Monte Carlo simulations. Section 6 outlines the labor market application and the empirical results. Section 7 concludes and discusses other possible applications. Appendices A-C gather some proofs and technical details while supplementary material is available in a Supplementary Appendix (SA henceforth).

## 2 The Environment and the Endogeneity Problem

We consider a (small) collection of  $F \geq 2$  firms, indexed by  $f = 1, \dots, F$ , bidding for the services of a (large) collection of workers, indexed by  $i = 1, \dots, N$ . According to the terminology in the introduction, firms (resp. workers) correspond to the principals (resp. agents). The workers are auctioned separately. In our application, the firms are the teams of the Indian Premier League (IPL) and the workers are the cricket players hired by these teams through English auctions. An observable reservation wage, denoted by  $W_i^r \geq 0$ , is attached to each worker  $i$  before this worker comes up for auction: it corresponds to the wage below which the worker cannot be sold. The winner of this auction (if there is one) is denoted  $f_i^w$ , and the corresponding (observable) final auction price determines the worker's wage 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 for 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, if there is no incumbent, we let  $f_i^{inc} = 0$ .

Before the auction for worker  $i$  starts, each firm  $f$  receives a one-dimensional private signal, denoted by  $s_{i,f}^{P,ms} \in \mathbb{R}$ . This signal is only observed by  $f$  and summarizes match-specific attributes regarding the worker's future post-auction performance if matched with  $f$ . There is also another one-dimensional signal associated to  $i$ , denoted by  $s_i^{P,ws} \in \mathbb{R}$ , which captures attributes relative to  $i$ 's future post-auction performance that are commonly valued by all firms. This signal always exists but is observed only by the incumbent  $f_i^{inc}$  if there is any (it is revealed to the incumbent through earlier interaction with this worker). Since in our empirical example workers' outcomes correspond to measures of performance, we shall call these signals the *productivity signals*. Furthermore,  $s_{i,f}^{P,ms}$  will be referred to as a *match-specific productivity signal* and  $s_i^{P,ws}$  as the *worker-specific productivity signal*.

The post-auction performance of worker  $i$  if matched with firm  $f$ , denoted  $y_{i,f} \in \mathbb{R}$ , is

---

Their setups are not, however, easily generalized to solve the additional econometric difficulty we face, namely endogenous matching.

assumed to take the following separable form:

$$y_{i,f} = \beta_f + \beta_x \cdot \mathbf{x}_{i,f} + \beta_{inc} \cdot INC_{i,f} + \tau(w_i) + \underbrace{s_{i,f}^{P,ms} + s_i^{P,ws}}_{:=u_{i,f}} + \epsilon_{i,f} \quad (1)$$

where the function  $\tau : [0, \infty) \rightarrow \mathbb{R} \cup \{-\infty\}$  captures the incentive effect. We make the normalization  $\tau(1) = 0$ . As outlined in the introduction,  $\tau(\cdot)$  is expected to be an increasing function either because better paid workers have greater incentives to exert more effort or because of reciprocity effects. The parameter  $\beta_f$  represents a firm-specific fixed effect,  $INC_{i,f}$  the dummy variable that is equal to one if  $f = f_i^{inc}$  and zero otherwise,  $\beta_{inc}$  the parameter measuring the effect of the incumbency indicator on post-auction performance,  $\mathbf{x}_{i,f}$  a vector of characteristics associated to the match  $(i, f)$  that are publicly observable before the auction starts,  $\beta_x$  a vector of parameters measuring the effects of the corresponding characteristics (all vectors are in bold).<sup>9</sup> The productivity signals  $s_i^{P,ws}$  and  $s_{i,f}^{P,ms}$  are normalized such that they are centered around zero conditional on the characteristics  $\mathbf{x}_{i,f}$  ( $f = 1, \dots, F$ ). Finally,  $\epsilon_{i,f}$  captures other post-auction performance determinants (including luck) that are unobserved by the firms at the bidding stage. We assume that the vectors  $(\epsilon_{i,1}, \dots, \epsilon_{i,F})$  are drawn independently across  $i$ . Furthermore, for each  $i$ , all  $F$  performance shocks are independent from all other variables in our model. The econometrician neither observes the productivity signals nor the performance shocks. From his/her perspective the error term in the post-auction performance equation is thus the sum of these three components, denoted  $u_{i,f}$ .

To estimate the post-auction performance equation (henceforth called shortly the performance equation), we face three kinds of problems. The first is that estimation is based on a selected sample. The selection arises because the wage and the post-auction performance of a given worker  $i$  are only observed if at least one firm decides to submit a bid above the reservation wage  $W_i^r$ . Intuitively, when there is no incumbent for a given worker  $i$ , the presence of this worker in the performance sample suggests that the match-specific productivity signal received by the winner is sufficiently high (bids are increasing in the productivity signals as formally shown in Section 3). Bearing in mind how  $u_{i,f}$  is defined in terms of this signal, the expectation of this error term is then likely to be higher for a worker who is in the sample than for one who is not.

Second, and even more crucially, we face the problem that workers are matched to firms in a non-random way. Indeed, if there is bidding at or above the reservation wage, only recorded is the post-auction performance  $y_{i,f_i^w}$ , while  $y_{i,f}$  is naturally unknown and counterfactual for all  $f \neq f_i^w$ . But the winner  $f_i^w$  is not an arbitrary firm. Instead, this is the firm whose bidding behavior was the most aggressive during the auction for worker  $i$ . This in turn implies that the winner's expected error term (conditional on worker  $i$  being sold, and on all observable variables in the performance equation),  $\mathbb{E}[u_{i,f} | w_i, \mathbf{x}_{i,f}, INC_{i,f}, f = f_i^w, w_i \geq W_i^r]$ , differs from the analogue expected error term for non-winners ( $f \neq f_i^w$ ) (the former is likely to be larger

---

<sup>9</sup>Note that  $\mathbf{x}_{i,f}$  may include contract characteristics established prior to the auction in which case the associated elements in  $\beta_x$  could represent other incentive effects of interest as well.



than the latter *ceteris paribus*). The matching of workers to firms is thus not random but is partly determined by the productivity signals, and hence the error term appearing in (1).

The final problem is related to the previous ones, but is nonetheless distinct. It concerns the fact that both  $w_i$  and  $INC_{i,f_i^w}$  are expected to be endogenous in (1). Large values of  $u_{i,f}$ , for  $f = 1, \dots, F$ , most likely indicate large values of  $s_i^{P,ws}$  and  $s_{i,f}^{P,ms}$ , for  $f = 1, \dots, F$ , which in turn should lead to a higher wage  $w_i$ . We thus expect the winner's expected error term  $\mathbb{E}[u_{i,f}|w_i, \mathbf{x}_{i,f}, INC_{i,f}, f = f_i^w, w_i \geq W_i^r]$  to be positively related to  $w_i$ . If we specify the incentive effect for example as  $\tau(w) = \tau \log(w)$ , a strong positive OLS estimate of  $\tau$  could then merely be an artefact caused by endogeneity bias. Similarly, the incumbent firm is expected to win more often when it receives a large worker-specific productivity signal, such that we expect the OLS estimator of  $\beta_{inc}$  to be upward biased.

Estimating the performance equation (1) using standard methods that ignore these problems of selection, non-random matching, and endogeneity of wages and incumbency status, would lead to biased estimates. Our Monte Carlo simulations confirm that the naive OLS estimates tend to be far away from the true parameter values. Our econometric strategy consists then in augmenting the performance equation with control terms that are explicitly micro-founded by an auction model. In a nutshell the approach works as follows: Let  $\mathcal{S}_i$  denote the set of variables observed by the econometrician just after the auction for worker  $i$  ended. Relying on and extending techniques from the structural econometrics of auctions, we model and estimate the expectations  $\mathbb{E}[s_{i,f_i^w}^{P,ms}|\mathcal{S}_i]$  and  $\mathbb{E}[s_i^{P,ws}|\mathcal{S}_i]$  up to a set of parameters, and then add the corresponding estimates in the performance equation.

The set  $\mathcal{S}_i$  includes the variables  $\mathbf{x}_{i,f}$ ,  $W_i^r$ ,  $INC_{i,f}$ , for all  $f$ , and the pair  $(w_i, f_i^w)$  for each worker  $i$  actually sold, but excludes the performance data. The set  $\mathcal{S}_i$  also includes  $\mathbf{z}_{i,f}$  (for all  $f$ ) which denotes a vector of variables observed by all bidders right before the auction for worker  $i$  starts. These variables, referred to as “auction shifters”, are assumed to affect how  $f$  values the services of  $i$ , but, crucially, they do not have a direct effect on  $i$ 's performance. While the identification of our model does not hinge on the availability of such variables, they will be helpful from an empirical perspective by playing the role of exogenous shifters, thereby reducing the collinearity between our control terms and the explanatory variables appearing in the performance equation. In Section 6, where we analyze sequential auctions for sport players,  $\mathbf{z}_{i,f}$  includes, among others, the remaining budget available to team  $f$  to purchase players and the characteristics of players already bought by this team once  $i$  is being auctioned. We will explain there why it is credible to assume that our chosen auction shifters can be excluded from the performance equation.

### 3 Matching through Auctions

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 post-auction performance. After having properly defined the auction rules and firms' payoff functions, we

characterize the equilibrium outcome as a function of the full set of signals.<sup>10</sup>

**Auction rules** Each worker is sold through an English auction with a public reservation wage. We formalize the game with the English button auction model where bidders observe when their opponents drop-out from the auction (Milgrom and Weber (1982)). For worker  $i$ , the incumbent (if any) first decides whether to enter the auction at the reservation wage  $W_i^r$ . After having observed whether the incumbent has entered or not the auction, non-incumbents then choose whether to enter the auction. There are now three possibilities. First, if there are no entrants at all, the auction stops and the worker remains definitely unmatched. Second, if there is a single entrant, the auction also stops but the worker is matched with this entrant and gets paid the reservation wage. Finally, if there are multiple entrants, the auction clock starts ticking at  $W_i^r$ , and moves up continuously. As the price goes up, entrants 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  $w_i$  paid by the winning bidder  $f_i^w$  to worker  $i$  corresponds to the auction termination/final price.<sup>11</sup> The pair  $(w_i, f_i^w)$  is referred to as the auction outcome. Throughout the paper, we call the bidding history of a firm (for a given worker) whether it has entered the auction or not and the wage at which it has dropped-out when it is not the winning firm. The bidding history is the collection of the bidding histories for each firm.

**Firms' payoff functions** We assume that firms are risk-neutral and hence maximize their expected payoff. The payoff derived by firm  $f$  from losing the auction for worker  $i$  is normalized to zero, while the payoff derived from winning the auction at wage  $w$  is assumed to take the form

$$V_{i,f} - w = \mathbb{E}_{\epsilon_{i,f}} [e^{\lambda \cdot [\overline{\mathcal{V}}(\mathbf{x}_{i,f}, \mathbf{z}_{i,f}, INC_{i,f}) + s_{i,f}^{I,ms} + s_i^{I,ws} + y_{i,f}]}] - w \quad (2)$$

where  $V_{i,f}$  designates how  $f$  values  $i$  as a function of  $w$ , the full set of signals, and the observable covariates associated to firm  $f$ , and  $\overline{\mathcal{V}}$  is a function of these covariates. By replacing  $y_{i,f}$  by the expression in (1) and taking the expectation of  $V_{i,f}$  with regards to the ex post shock  $\epsilon_{i,f}$ , we obtain then that

$$V_{i,f} = e^{\lambda \cdot [\overline{V}_{i,f} + s_{i,f}^{I,ms} + s_i^{I,ws} + s_{i,f}^{P,ms} + s_i^{P,ws} + \tau(w)]} \quad (3)$$

with  $\overline{V}_{i,f} := \overline{\mathcal{V}}(\mathbf{x}_{i,f}, \mathbf{z}_{i,f}, INC_{i,f}) + \beta_f + \beta_x \cdot \mathbf{x}_{i,f} + \beta_{inc} \cdot INC_{i,f} + (\ln(\mathbb{E}_{\epsilon_{i,f}} [e^{\lambda \cdot \epsilon_{i,f}}]))/\lambda$ . The additional signals  $s_{i,f}^{I,ms} \in \mathbb{R}$  (for  $f = 1, \dots, F$ ) and  $s_i^{I,ws} \in \mathbb{R}$  are one-dimensional and referred to as the *intrinsic signals*: they are intrinsic to the firms insofar as they enter their payoff function but do not directly affect the worker's performance. The intrinsic signals reflect that firms may value workers for other motives besides the individual performance measure. These signals may for instance measure other individual performance measures that are unobserved to the

<sup>10</sup>We abstract from bid increments which are often implemented by auctioneers in English auctions. To deal with increments, Haile and Tamer (2003) specify the auction model only partially, i.e., adopt an incomplete model approach. Such an approach would prevent us to identify the expectations  $\mathbb{E}[s_{i,f}^{P,ms} | \mathcal{G}_i]$  and  $\mathbb{E}[s_{i,f}^{P,ws} | \mathcal{G}_i]$ .

<sup>11</sup>The auction-determined wage  $w_i$  is the sole source of revenue for worker  $i$ . In particular, there is no performance-based remuneration. Although it is beyond the scope of this paper to account for such type of payment structure, our methodology could be extended to contingent auctions.

econometrician. They could also capture attributes that affect the global workplace conditions of the firm where  $i$  is hired. The information structure concerning these additional signals is analogous to the one adopted for the productivity signals: for any given worker  $i$ , each firm  $f$  is assumed to receive the match-specific signal  $s_{i,f}^{I,ms}$  and the incumbent (if any) receives in addition the worker-specific signal  $s_i^{I,ws}$ . Like the worker-specific productivity signal  $s_i^{P,ws}$ ,  $s_i^{I,ws}$  reflects insider information held by the incumbent thanks to the previous relationship with the worker. This signal always exists, even if there is no incumbent for worker  $i$ .

Since the function  $\bar{\mathcal{V}}$  depends on  $\mathbf{x}_{i,f}$  and  $INC_{i,f}$ , these variables have a direct effect on how  $f$  values  $i$  and an indirect effect through  $y_{i,f}$ . This function also depends on  $\mathbf{z}_{i,f}$ , so the auction shifters only have a direct effect because they are excluded from the performance equation. The exponential is used to guarantee that a firm's payoff from winning at wage zero is always positive. Firm  $f$  takes the expectation with respect to  $\epsilon_{i,f}$  because at the bidding stage this productivity shock is unknown; the signals and other variables can be taken outside the expectation operator because they are independent of the shock. The parameter  $\lambda > 0$  measures to which extent the post-auction performance  $y_{i,f}$  matters in valuing agents in monetary terms.<sup>12</sup>

We wish to make a few comments on the payoff function. 1) Because of the incentive effect  $\tau(w)$ , our model departs from the quasi-linear paradigm generally adopted in both the theoretical and empirical auction literature where the payoff of the winning bidder is assumed to be linear in the auction price. 2) Our model involves interdependent values since the payoff of each firm depends on the worker-specific signals,  $s_i^{I,ws}$  and  $s_i^{P,ws}$ , which are observed only by the incumbent. 3) Firm  $f$ 's preferences depend on its own matching-specific signals and on the worker-specific signals through their sum: this is a kind of separability assumption that is crucial to deal with environments with multi-dimensional signals both from a theoretical and econometric point of view.

**Assumptions** Before stating the assumptions that are needed, we introduce some additional notations. For each worker  $i$ , we let  $\mathbf{s}_{i,f}$  denote the vector of firm  $f$ 's private signals, i.e.,  $\mathbf{s}_{i,f}^{inc} = (s_{i,f}^{I,ms}, s_{i,f}^{P,ms}, s_i^{I,ws}, s_i^{P,ws})$  for the incumbent (if any) and  $\mathbf{s}_{i,f} = (s_{i,f}^{I,ms}, s_{i,f}^{P,ms})$  otherwise. We also let  $\bar{\mathbf{s}}_{i,f}$  denote the sum of  $f$ 's private signals which is referred to as  $f$ 's aggregate signal. For the incumbent, we have  $\bar{\mathbf{s}}_{i,f}^{inc} = s_{i,f}^{I,ms} + s_{i,f}^{P,ms} + s_i^{I,ws} + s_i^{P,ws}$ , and otherwise  $\bar{\mathbf{s}}_{i,f} = s_{i,f}^{I,ms} + s_{i,f}^{P,ms}$ . Throughout our analysis, we impose the following assumptions on the signals:

**A1:** i) The vector  $((s_{i,f}^{I,ms}, s_{i,f}^{P,ms}, \mathbf{x}_{i,f}, \mathbf{z}_{i,f})_{f=1,\dots,F}, s_i^{I,ws}, s_i^{P,ws}, f_i^{inc})$  is i.i.d. across  $i$ . ii) For each  $i$ , the  $F+1$  pairs of signals  $(s_{i,f}^{I,ms}, s_{i,f}^{P,ms})_{f=1,\dots,F}$  and  $(s_i^{I,ws}, s_i^{P,ws})$  are jointly independent, conditional on  $(\mathbf{x}_{i,f}, \mathbf{z}_{i,f})_{f=1,\dots,F}$  and  $f_i^{inc}$ . iii) For each  $i$  and  $f = 1, \dots, F$ , each signal belonging to  $\mathbf{s}_{i,f}$  has a mean equal to zero conditional on  $(\mathbf{x}_{i,f}, \mathbf{z}_{i,f})_{f=1,\dots,F}$  and  $f_i^{inc}$ .

In the empirical auction literature, A1 i) is standard and guarantees that our estimators have the usual asymptotic properties. The independence restrictions in A1 ii) are crucial for two kinds of reasons: On the hand, they imply the independence of the aggregate signals

<sup>12</sup>The parameter  $\lambda$  also multiplies the intrinsic signals and the function  $\bar{\mathcal{V}}(\cdot)$  in (2) but this is without loss of generality once the intrinsic signals and the function  $\bar{\mathcal{V}}(\cdot)$  are properly normalized up to multiplicative constants.

$\bar{s}_{i,f}$  across all  $f$  which allows to identify non-parametrically the individual bid distributions from the winning bid distribution in English auctions (Athey and Haile (2002)). On the other hand, these restrictions imply that the sum of worker-specific signals  $s_i^{I,ws} + s_i^{P,ws}$  is independent of the sum of the incumbent's match-specific signals  $s_{i,f}^{I,ms} + s_{i,f}^{P,ms}$ , which allows us to use a deconvolution argument as a key step to obtain our non-parametric identification result. A1 ii) coupled with the fact that signals enter the payoff function in an additive manner is reminiscent of the restrictions used to deal with unobserved heterogeneity (see Perrigne and Vuong (2021)). A1 i) and ii) are also crucial from a theoretical point of view: i) ensures that each auction can be analyzed separately from the others, i.e., that the equilibrium strategy of firm  $f$  for worker  $i$  is unrelated to the signals regarding workers  $i' \neq i$ ; ii) guarantees that all non-incumbents always share the same beliefs about the distribution of both worker-specific signals (whether there is an incumbent or not). A1 iii) is without loss of generality to identify causally the incentive effect  $\tau(\cdot)$  (if the signals were not mean-independent of the observable covariates, we could renormalize them by putting their mean in the term  $\bar{V}_{i,f}$ ).<sup>13</sup> We emphasize that A1 does not impose the pair of match-specific signals  $(s_{i,f}^{I,ms}, s_{i,f}^{P,ms})$  to be identically distributed across firms, nor does it restrict intrinsic and productivity signals to be independent. Thus,  $s_{i,f}^{I,ms}$  and  $s_{i,f}^{P,ms}$  may be correlated variables, and  $s_i^{I,ws}$  and  $s_i^{P,ws}$  as well.<sup>14</sup> The following assumption concerns the incentive function  $\tau(w)$ .

**A2:** The function  $\psi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ ,<sup>15</sup> defined by  $\psi(w) := w \cdot e^{-\lambda\tau(w)}$ , is an increasing bijection.

This assumption guarantees that if a firm is prepared to pay a price  $w$  for a worker then this firm would be prepared to pay any wage below  $w$ . When  $\tau(w) = \tau \cdot \log(w)$ , the specification chosen in our application, then A2 is equivalent to  $\lambda \cdot \tau < 1$ .

Before turning to the equilibrium analysis, we introduce the notion of *valuation*, i.e., the maximum wage that  $f$  would be prepared to bid under complete information regarding the signals. If  $f$  could observe the worker-specific signals  $s_i^{P,ws}$  and  $s_i^{I,ws}$  in addition to its match-specific signals, then thanks to A2 this firm's payoff function crosses zero only once, at  $w_{i,f}^* \equiv \psi^{-1}(e^{\lambda \cdot [\bar{V}_{i,f} + s_{i,f}^{I,ms} + s_i^{I,ws} + s_{i,f}^{P,ms} + s_i^{P,ws}]})$ . At this cutoff wage,  $f$  is indifferent between winning or losing the auction for worker  $i$ , and its payoff would be strictly positive (resp. negative) if the wage is below (resp. above)  $w_{i,f}^*$ . This cutoff wage is referred to next as firm  $f$ 's valuation for worker  $i$ .

**Equilibrium analysis** We report below informally the fundamental properties of the equilibrium that sustain our econometric methodology. The formal analysis is relegated to Appendix A, in particular in Proposition A.1. Therein we show that, under A1, A2 and some technical assumptions (TA1-TA3), there is a unique perfect Bayesian equilibrium. The equilibrium strategies are described in a detailed manner since the Bayesian updating, required to compute our control function terms, relies on the exact expression of these strategies.

<sup>13</sup>Once A1 ii) is strengthened by assuming that signals' distributions do not depend on the variables  $INC_{i,f}$  (as we do later in A3), then A1 iii) is also without loss of generality to identify causally the parameter  $\beta_{inc}$ . A1 iii) is required, however, if one wishes to give a causal interpretation to the estimate of  $\beta_x$  as well.

<sup>14</sup>For instance, a positive correlation between the latter two signals could reflect that the best workers in terms of individual performance are also those that contribute most to a favorable workplace environment.

<sup>15</sup>We use the convention that  $\mathbb{R}_+ = (0, +\infty)$ .

Since all  $N$  auctions can be analyzed independently thanks to A1, we can arbitrarily focus on one of them, say the auction for worker  $i$ . The payoff function of the incumbent (if any) depends on the vector of signals  $\{s_{i,f}\}_{f=1,\dots,F}$  only through its own signal  $s_{i,f_i^{inc}}$ . From the incumbent's perspective, the auction game is a private value English auction and, as usual, it is then a dominant strategy to bid until its valuation: The equilibrium strategy of the incumbent consists in not entering the auction if  $w_{i,f_i^{inc}}^*$  is strictly below  $W_i^r$ , and otherwise remain active until the clock reaches its valuation. Note that the incumbent's strategy depends on its signal solely through the aggregate signal  $\bar{s}_{i,f_i^{inc}}$ .

The equilibrium strategies of non-incumbents are slightly more subtle because they do not know their valuation. Since the payoff function of  $f$  for worker  $i$  depends on the vector of match- and worker-specific signals only through  $\bar{s}_{i,f}$  and  $s_i^{I,ws} + s_i^{P,ws}$ , this firm's bidding incentives depend on its own aggregate signal and its beliefs regarding the sum of worker-specific signals. Thanks to A1 all non-incumbents share the same beliefs, which depend on whether an incumbent is present or not, and, if present, on the bidding history of  $f_i^{inc}$ .

In the absence of an incumbent, the cutoff wage that makes a non-incumbent firm  $f$  indifferent between winning or losing the auction is  $\psi^{-1}\left(e^{\lambda \cdot [\bar{V}_{i,f} + \bar{s}_{i,f}]} \cdot \mathbb{E}[e^{\lambda \cdot (s_i^{I,ws} + s_i^{P,ws})}]\right)$ , where the expectation is with respect to  $s_i^{I,ws} + s_i^{P,ws}$ .<sup>16</sup> The equilibrium strategy for  $f$  is to enter the auction if the cutoff wage exceeds  $W_i^r$ , and then remain active until the clock has reached the cutoff wage. Suppose now instead that there is an incumbent and that the current value of the auction clock is  $\tilde{w} \geq W_i^r$ . If  $f_i^{inc}$  dropped out of the auction somewhere before this value was reached while at least two non-incumbents are still active, then the equilibrium strategy is similar as just before except that the expectation appearing in the cutoff wage is now conditional on the incumbent's bidding history.<sup>17</sup> To complete the characterization of the equilibrium strategies, there is a last remaining case, namely the one where  $f_i^{inc}$  is still active at  $\tilde{w}$  while there is at least one active non-incumbent. As detailed in Appendix A, the analysis is then similar to [Milgrom and Weber \(1982\)](#): a non-incumbent should remain active if and only if its expected payoff at  $\tilde{w}$  would be positive under the belief that  $f_i^{inc}$  drops out immediately at this clock value, that is to say under the belief that  $w_{i,f_i^{inc}}^* = \tilde{w}$ .

One of the main take-aways from the equilibrium analysis, and which will be used throughout our econometric analysis, is that the equilibrium is unique and involves pure strategies that are continuously increasing in firms' aggregate signals. Those are key properties that need to be derived to adapt our approach to other auction formats. Note that auctions with interdependent values and multidimensional signals are known to raise important issues in terms of equilibrium existence ([Jackson \(2009\)](#)) or equilibrium multiplicity ([Heumann \(2019\)](#)). Our model where there is at most one bidder –the incumbent– who is informed about commonly valued elements is a way to tackle these difficult issues while keeping a high degree of generality regarding bidders' signal distributions.<sup>18</sup> For our econometric analysis, another important

<sup>16</sup>For simplicity we omit here that expectations with respect to signals are also conditional on  $(\mathbf{x}_{i,f}, \mathbf{z}_{i,f})_{f=1,\dots,F}$ .

<sup>17</sup>Each remaining non-incumbent thus calculates the expectation given  $w_{i,f_i^{inc}}^* \leq W_i^r$  (if the incumbent did not enter the auction at all) or given  $w_{i,f_i^{inc}}^*$  (if the incumbent quit somewhere between  $W_i^r$  and  $\tilde{w}$ ).

<sup>18</sup>In contrast, [Goeree and Offerman \(2003\)](#) and [Heumann \(2019\)](#) analyze models where all bidders receive

property is Corollary 1 (in Appendix A). It states the model primitives that need to be identified in order to characterize the equilibrium and the distribution of the bidding history.

**Related models in the literature** Our model is reminiscent of [Goeree and Offerman \(2003\)](#) and [Heumann \(2019\)](#)'s models with bi-dimensional signals, one capturing privately valued attributes and another commonly valued attributes. Their equilibrium characterization also relies on projecting bidders' multidimensional signals onto a one-dimensional aggregate signal.<sup>19</sup> In contrast to these papers, our model departs from the quasi-linear paradigm. Furthermore, we consider that at most one bidder can receive commonly valued signals. Our auction model is also related to models wherein a single bidder receives insider information. [Engelbrecht-Wiggans et al. \(1983\)](#)'s analysis of the first price auction is the seminal contribution, later extended by [Hendricks et al. \(1994\)](#) when the seller uses a secret reserve price. However, both consider pure common value models while our setup also involves idiosyncratic private signals. Finally, our paper is related to papers in the labor literature which consider auction models with an incumbent (see [Greenwald \(1986\)](#) and [Lazear \(1986\)](#)). Such models have been used to rationalize the lifetime dynamics of wages ([Pinkston \(2009\)](#)) or firms' incentives to train their employees ([Acemoglu and Pischke \(1998\)](#)).

## 4 The Econometric Methodology

This section is devoted to the econometric aspects of the paper. Section 4.1 presents the first step of our econometric strategy: the non-parametric identification of the bidding model from the auction data. Section 4.2 presents then the second step of our strategy: the identification of the performance equation through control functions that are identified from the primitives of the auction model. Section 4.3 presents our estimation method in the parametric setup we use in both our Monte Carlo simulations and our empirical application. Section 4.4 discusses how to adapt our methodology when the auction data never involves an incumbent. Finally, discussions on related literature, alternative identification strategies and extensions to other auction formats are relegated to the SA.

### 4.1 Identification of the auction model

We assume throughout this section that the reservation wage is always equal to zero:  $W_i^r = 0$  for all  $i$ . We discuss in the SA how identification can be achieved when this assumption is relaxed. To obtain our identification result, we need a few additional assumptions, and some new notations are required. Let  $COV_i := (x_{i,f}, z_{i,f})_{f=1,\dots,F}$  be the vector of observable covariates associated with worker  $i$ . According to A1, the distribution of the vector of private

---

private information about commonly valued elements. To guarantee equilibrium existence, [Goeree and Offerman \(2003\)](#) consider a fully symmetric model while [Heumann \(2019\)](#) imposes a Gaussian information structure.

<sup>19</sup>Another motive for using such aggregation techniques is when the action space of each bidder is multidimensional as in scoring auctions ([Asker and Cantillon \(2008\)](#)), where the equilibrium scores (which determine the auction outcome) are functions of bidders' aggregate signals.

signals  $\{s_{i,f}\}_{f=1,\dots,F}$  depends on  $i$  only through  $COV_i$  and the identity of the incumbent  $f_i^{inc}$ . A3 below assumes in addition that these signal distributions do not depend on whether an incumbent is present, nor on the incumbent's identity if there is one.

**A3:** The distribution of the match-specific signals  $(s_{i,f}^{I,ms}, s_{i,f}^{P,ms})$  (for each  $f = 1, \dots, F$ ) and the distribution of the worker-specific signals  $(s_i^{I,ws}, s_i^{P,ws})$  only depend on  $COV_i$ .

A3 is our main exclusion restriction for identification: the private signals are distributed identically in the subsets with and without an incumbent.

Let  $G^{ws}(\cdot|COV_i)$  denote the distribution of  $\lambda \cdot (s_i^{I,ws} + s_i^{P,ws})$  and  $G_f^{ms}(\cdot|COV_i)$  (for all  $f$ ) the distribution of  $\lambda \cdot (s_{i,f}^{I,ms} + s_{i,f}^{P,ms})$ , conditional on  $COV_i$ .<sup>20</sup> From our equilibrium analysis (see Corollary 1 in Appendix A), the strategy of each firm  $f \in \{1, \dots, F\}$ , in the auction for worker  $i$ , depends solely on  $\lambda \cdot \bar{s}_{i,f}$ , the identity of the incumbent  $f_i^{inc}$  and the following primitives: the distributions  $\{G_f^{ms}(\cdot|COV_i)\}_{f=1,\dots,F}$  and  $G^{ws}(\cdot|COV_i)$ , the scalars  $\lambda \cdot \bar{V}_{i,f}$  for each  $f = 1, \dots, F$ , and the function  $\psi$ . In order to identify the bidding model, these are thus the primitives of interest we wish to identify. Since the scalars  $\bar{V}_{i,f}$ ,  $f = 1, \dots, F$ , depend on  $i$  only through  $COV_i$  and  $INC_{i,f}$ ,  $f = 1, \dots, F$ , we get on the whole that the distribution of all cutoff wages and hence the distribution of the bidding history depend on worker  $i$  only through  $COV_i$  and  $f_i^{inc}$ . This is important for the Bayesian updating exercises in Section 4.2, but also because it implies that the bid distributions defined right below need not be indexed by  $i$ .

Let  $H_f^{inc}(\cdot|COV, f^{inc})$  denote the distribution of firm  $f$ 's valuation for worker  $i$ ,  $w_{i,f}^*$ , if  $COV_i = COV$  and  $f_i^{inc} = f^{inc}$ . For  $f = f^{inc}$ , this corresponds to the incumbent's bid distribution. Let  $H_f^{NI}(\cdot|COV)$  denote the bid distribution of firm  $f$  for worker  $i$  given  $COV_i = COV$  and when the auction involves no incumbent, that is to say the distribution of the cutoff wage

$$w_{i,f}[NI](\bar{s}_{i,f}) := \psi^{-1} \left( e^{\lambda \cdot [\bar{V}_{i,f} + \bar{s}_{i,f}]} \cdot \mathbb{E}[e^{\lambda \cdot (s_i^{I,ws} + s_i^{P,ws})} | COV_i = COV] \right). \quad (4)$$

Similarly, let  $H_f^{ACT}(\cdot|COV, f^{inc})$  denote the distribution of the cutoff wage up to which  $f \neq f_i^{inc}$  remains active in equilibrium when the incumbent  $f_i^{inc} = f^{inc}$  is still active. This cutoff wage is characterized (see Appendix A for details) as the unique solution, in  $u$ , of the equation

$$e^{\lambda \cdot [\bar{V}_{i,f} + \bar{s}_{i,f}]} \cdot \mathbb{E}[e^{\lambda \cdot [s_i^{I,ws} + s_i^{P,ws}]} | w_{i,f_i^{inc}}^* = u, COV_i = COV, f_i^{inc} = f^{inc}] = \psi(u). \quad (5)$$

Next we define  $S_i \subseteq \{1, \dots, F\}$ , the publicly observable subset of firms that are potential participants in the auction for worker  $i$ .<sup>21</sup> In our equilibrium analysis in Section 3, we have 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 now say that the auction for worker  $i$  involves an incumbent if  $f_i^{inc} \in S_i$ . Contrary to the distribution of the auction outcome  $(w_i, f_i^w)$ , the distri-

<sup>20</sup>A1 and A3 imply that these conditional distributions do not depend on worker  $i$  (apart from  $COV_i$ ), which is why they are not indexed by  $i$ .

<sup>21</sup>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.

butions of the underlying cutoff wages  $H_f^{inc}(\cdot|COV, f^{inc})$ ,  $H_f^{NI}(\cdot|COV)$  and  $H_f^{ACT}(\cdot|COV, f^{inc})$  (for  $f^{inc}, f \in S$ ), do not depend on the set of potential participants  $S_i$ .

The following assumption defines the set of variables observed by the econometrician.

**A4:** For each  $i$ , the information set  $\mathcal{S}_i$  contains at least the following information: the covariates  $COV_i$ , the incumbency indicator  $f_i^{inc}$ , the set of potential participants  $S_i$ , and the auction outcome  $(w_i, f_i^w)$ .

Note that  $\mathcal{S}_i$  contains the explanatory variables appearing in the performance equation (1) not only for the winning bidder but also for the losing bidders. Note also that we do not assume here that the econometrician observes the full bidding history, i.e., the entry and drop-out decisions of all potential participants: The econometrician thus may know less than the firms since the latter are assumed to observe the bidding behavior of all competitors and in particular of the incumbent.

The two following assumptions concern the distribution of  $S_i$ , and the link between the function  $\psi$  and the distributions  $H_f^{ACT}(\cdot|COV, f^{inc})$ . From now on  $\mathcal{S}_{COV}$  denotes the support of the distribution of  $COV$ .

**A5:** Conditional on  $\{COV_i, f_i^{inc}\}_{i=1, \dots, N}$ , the sets of potential participants  $(S_1, \dots, S_N)$  are drawn independently from the private signals, and they are drawn independently of each other.

**A6:** For any pair  $\tilde{f}, \tilde{f}' \in \{1, \dots, F\}$ , and given the distributions  $H_f^{NI}(\cdot|COV)$  (for  $f = 1, \dots, F$ ) and  $H_{\tilde{f}}^{inc}(\cdot|COV, \tilde{f})$  for any  $COV \in \mathcal{S}_{COV}$ , the function that maps  $\psi$  (from the set of increasing bijections on  $\mathbb{R}_+$  with  $\psi(1) = 1$ ) into the set of distributions  $\{H_{\tilde{f}'}^{ACT}(\cdot|COV, \tilde{f}')\}_{COV \in \mathcal{S}_{COV}}$  is injective.

A5 guarantees that the set of participants  $S_i$  is determined exogenously conditional on  $COV_i$  and  $f_i^{inc}$  and is thus a complement to the independence restrictions in A1. How  $S_i$  varies (or not) across  $i$  appears through assumption TA4 in Appendix B. We emphasize that if  $F = 2$ , then the set of assumptions we use for identifying the auction model is consistent with  $S_i = \{1, 2\}$  for each  $i$ , i.e., identification does not rely on exogenous variations in the set of potential participants. Whether A6 is satisfied is not easy to verify in general because the function  $\psi$  appears in (5) not only explicitly but also implicitly through the term  $w_{i, f^{inc}}^*$ . Nevertheless, we do think that this identification restriction is not very strong. In particular, under the linear-log parametric restriction  $\tau(w) = \tau \cdot \log(w)$  and assumptions A1-A5 and TA1-TA3, we show in Appendix B that it will always be satisfied and this even if the set  $\mathcal{S}_{COV}$  is a singleton. The logic of the proof is instructive: the parameter  $\theta := \tau \cdot \lambda$  (and hence  $\psi(w) = w^{1-\theta}$ ) is identified by comparing the  $q$ -th bid quantile of  $H_f^{NI}(\cdot|COV)$  and  $H_{\tilde{f}}^{ACT}(\cdot|COV, \tilde{f})$  for an arbitrary quantile  $q$ , an arbitrary set of covariates  $COV$  and an arbitrary firm  $f \neq \tilde{f}$ . Given that equilibrium bid functions are increasing in firms' aggregate signals and given A3, these bid-quantiles correspond to the same aggregate signal for  $f$ . Fixing  $f$ 's aggregate signal and conditional on  $COV$ , comparing its equilibrium bid when there is no incumbent with the bid submitted when this firm competes with one opponent, the incumbent  $\tilde{f}$ , allows then to identify  $\theta$ . Intuitively, this logic could be extended to the case where the function  $\tau(\cdot)$  includes multiple unknown parameters by exploiting multiple quantiles of these distributions. Furthermore, variations in the covariates is



also expected to help to satisfy A6.

Two final assumptions are needed for identification of our bidding model. They are technical restrictions and therefore relegated to Appendix B (TA4 and TA5). We are now ready to establish that the primitives of interest of our auction model are identified in the following proposition.

**Proposition 4.1.** *Assume A1-A6 and TA1-TA5. In the English auction without reservation wage, the function  $\psi$ , the distributions  $G^{ws}(\cdot|COV)$  and  $G_f^{ms}(\cdot|COV)$ ,  $f = 1, \dots, F$  for any covariate  $COV \in \mathcal{S}_{COV}$ , and the scalars  $\lambda \cdot \bar{V}_{i,f}$ ,  $f = 1, \dots, F$  (for any  $i$  such that  $COV_i \in \mathcal{S}_{COV}$  and for any realization of  $f_i^{inc}$ ), are identified.*

This is a non-parametric identification result in the sense that both the signal distributions  $G^{ws}(\cdot|COV)$  and  $G_f^{ms}(\cdot|COV)$ ,  $f = 1, \dots, F$ , and the function  $\psi$  are left unspecified. Note, however, that Proposition 4.1 does not state that  $\lambda$  is identified. Indeed, for two pairs  $(\lambda, \psi)$  and  $(\lambda', \psi')$  such that  $\lambda\tau(\cdot) = \lambda'\tau'(\cdot)$ , the implied auction outcomes would be observationally equivalent. In words, we can not disentangle from the bidding data alone the incentive effect from the parameter  $\lambda$ .

The proof of Proposition 4.1 is relegated to Appendix B. Here we just sketch the main ideas of the proof. The independence of the aggregate signals  $\bar{s}_{i,f}$  across firms (conditional on  $COV$ ) implies the independence of the equilibrium cutoff wages. Since we observe the second highest cutoff wage (which corresponds to the final auction price in the English auction) and the identity of the bidder having the highest cutoff wage (which corresponds to the auction winner), we can apply [Athey and Haile \(2002\)](#): for any  $COV \in \mathcal{S}_{COV}$ , the distributions  $\{H_f^{NI}(\cdot|COV)\}_{f=1, \dots, F}$  are identified from the sub-sample of workers  $i$  for which the set of potential participants does not include an incumbent (i.e.,  $f_i^{inc} \notin S_i$ ), and the distributions  $H_f^{inc}(\cdot|COV, f)$  and  $H_{f'}^{ACT}(\cdot|COV, f)$  are identified from the auctions  $i$  for which  $S_i = \{f, f'\}$  and  $f_i^{inc} = f$ . From A6,  $\psi$  is then also identified. Using the functions that map firms' aggregate signals to their cutoff wage, we identify, for each  $f$ ,  $G_f^{ms}(\cdot|COV)$  the (zero-mean) distribution of  $\lambda \cdot (s_{i,f}^{I,ms} + s_{i,f}^{P,ms})$  conditional on  $COV_i = COV$ , from the distribution  $H_f^{NI}(\cdot|COV)$ , and the (zero-mean) distribution of  $\lambda \cdot (s_{i,f}^{I,ms} + s_{i,f}^{P,ms} + s_i^{I,ws} + s_i^{P,ws})$ , conditional on  $COV_i = COV$ , from the distribution  $H_f^{inc}(\cdot|COV, f)$ . Since match-specific and worker-specific signals are assumed to be independent (Assumption A1 ii), and the distribution of  $\lambda \cdot (s_{i,f}^{I,ms} + s_{i,f}^{P,ms})$  is the same both in auctions where  $f$  is the incumbent and in auctions where there is no incumbent (Assumption A3), we get from a deconvolution argument that  $G^{ws}(\cdot|COV)$ , i.e., the distribution of  $\lambda \cdot (s_i^{I,ws} + s_i^{P,ws})$ , is identified as well. Finally, we identify the scalars  $\lambda \cdot \bar{V}_{i,f}$  for each  $i$  such that  $COV_i \in \mathcal{S}_{COV}$  both when  $f = f_i^{inc}$  and when  $f \neq f_i^{inc}$ : when  $f \neq f_i^{inc}$ , we identify  $\lambda \cdot \bar{V}_{i,f}$  from  $H_f^{NI}(\cdot|COV)$  and  $\psi$ ; when  $f = f_i^{inc}$  we identify  $\lambda \cdot \bar{V}_{i,f}$  from  $H_f^{inc}(\cdot|COV, f)$ ,  $\psi$  and the distribution  $G^{ws}(\cdot|COV)$ .

We emphasize that firms' equilibrium bids for worker  $i$  (as a function of their private signals) do not depend on  $S_i$  but rather on  $COV_i$ ,  $f_i^{inc}$  and whether  $f_i^{inc} \in S_i$ . Our identification result thus does not exploit exogenous variations of the set of potential participants (as in [Guerre](#)

et al. (2009)) but rather exogenous variations of the incumbency variable  $f_i^{inc}$  paired with the exclusion restriction A3.

## 4.2 Identification of the performance equation

The control function approach amounts to modeling the conditional expectation  $\mathbb{E}[u_{i,f_i^w} | \mathcal{I}_i] = \mathbb{E}[s_{i,f_i^w}^{P,ms} + s_{i,f_i^w}^{P,ws} | \mathcal{I}_i]$  using our auction model and bidding data. Since the distributions of  $s_{i,f}^{P,ms}$  and  $s_{i,f}^{P,ws}$  are not identified from our first stage (not even up to a multiplicative constant), we impose the following semiparametric restriction on the link between on the one hand the distributions of  $s_{i,f}^{P,ms}$  and  $s_{i,f}^{I,ms} + s_{i,f}^{P,ms}$  for each  $f$ , and on the other the distributions of  $s_{i,f}^{P,ws}$  and  $s_{i,f}^{I,ws} + s_{i,f}^{P,ws}$ .

**A7:** For each  $COV \in \mathcal{S}_{COV}$ , we have:  $\mathbb{E}[s_{i,f}^{P,ms} | s_{i,f}^{I,ms} + s_{i,f}^{P,ms} = x, COV_i = COV] = \sum_{l=1}^L \gamma_{l,f}^{ms}(COV) \cdot (\lambda \cdot x)^l$  and  $\mathbb{E}[s_{i,f}^{P,ws} | s_{i,f}^{I,ws} + s_{i,f}^{P,ws} = x, COV_i = COV] = \sum_{l=1}^L \gamma_{l,f}^{ws}(COV) \cdot (\lambda \cdot x)^l$ , for any  $x \in \mathbb{R}$  and each  $f \in \{1, \dots, F\}$ , and the functions  $x \rightarrow \sum_{l=1}^L \gamma_{l,f}^{ms}(COV) \cdot (\lambda \cdot x)^l$  ( $f = 1, \dots, F$ ) and  $x \rightarrow \sum_{l=1}^L \gamma_{l,f}^{ws}(COV) \cdot (\lambda \cdot x)^l$  are nondecreasing.

We do not impose any restriction on  $L$ , the order of the polynomial.<sup>22</sup> The scalar parameters  $\gamma_{l,f}^{ms}(COV)$ ,  $f = 1, \dots, F$ , and  $\gamma_{l,f}^{ws}(COV)$  depend non-parametrically on the vector of covariates  $COV$ . The monotonicity restriction in A7 is mild<sup>23</sup> and implies in particular that if a firm  $f$  is a non-incumbent and bids more conditional on the observable covariates, then the worker is expected to perform better if matched with  $f$  ceteris paribus. This restriction allows us to rely on one-sided tests in our application. Note that we do not put any restriction on the sign of the correlation between  $s_{i,f}^{I,ms}$  and  $s_{i,f}^{P,ms}$ , or between  $s_{i,f}^{I,ws}$  and  $s_{i,f}^{P,ws}$ .

Let us define the augmented performance equation:

$$y_{i,f_i^w} = \beta_{f_i^w}(COV_i) + \beta_{inc} \cdot INC_{i,f_i^w} + \frac{1}{\lambda} \cdot \log\left(\frac{w_i}{\psi(w_i)}\right) + \sum_{l=1}^L \gamma_{l,f_i^w}^{ms}(COV_i) \cdot CF_i^{ms}[l] + \sum_{l=1}^L \gamma_{l,f_i^w}^{ws}(COV_i) \cdot CF_i^{ws}[l] + \xi_{i,f_i^w} \quad (6)$$

where  $\beta_{f_i^w}(COV_i) := \beta_{f_i^w} + \beta_x \cdot \mathbf{x}_{i,f_i^w}$ , the match-specific control terms are  $CF_i^{ms}[l] := \mathbb{E}[[\lambda \cdot (s_{i,f_i^w}^{I,ms} + s_{i,f_i^w}^{P,ms})]^l | \mathcal{I}_i]$  (for  $l = 1, \dots, L$ ), and the worker-specific control terms are  $CF_i^{ws}[l] := \mathbb{E}[[\lambda \cdot (s_{i,f_i^w}^{I,ws} + s_{i,f_i^w}^{P,ws})]^l | \mathcal{I}_i]$  (for  $l = 1, \dots, L$ ). In Appendix C, we show that these control terms are identified from the auction data and that the error term  $\xi_{i,f_i^w}$  is equal to  $u_{i,f_i^w} - \mathbb{E}[u_{i,f_i^w} | \mathcal{I}_i]$  and hence is uncorrelated to all regressors appearing in (6). How to compute the control terms (which relies on tedious Bayesian updating exercises) is detailed in the SA.

Let us now discuss the identification of the augmented performance equation (6) conditional

<sup>22</sup>Instead of the univariate polynomials basis  $(x, x^2, \dots, x^L)$ , we could also use an alternative basis of functions. Following Komarova (2017)'s sieve estimators, Kong (2021) uses estimators based on the Bernstein polynomial basis which allows to impose easily monotonicity constraints on the function one wishes to estimate.

<sup>23</sup>This restriction was implicitly made in our informal discussion (in Section 2) regarding the bias of the OLS estimators of the incentive effect and the parameter  $\beta_{inc}$ .

on a given realization for the vector of covariates  $COV \in \mathcal{S}_{COV}$ . For any given set of participants  $S$  with  $|S| \geq 2$ , our full support assumption in TA2 guarantees that the support of the auction outcome  $(w_i, f_i^w)$  is  $[0, +\infty) \times S$ : any potential participant may win the auction at any strictly positive wage. Furthermore, given A5 (and also TA5 in Appendix B), all firms can be observed as a participant in an auction without an incumbent and one of them, denoted next  $\tilde{f}$ , is observed also as an incumbent.

If we assume that for any given  $f \in \{1, \dots, F\}$ , the random vector  $(1, \log(w_i/\psi(w_i)), CF_i^{ms}[1], \dots, CF_i^{ms}[L])$  is of full rank conditional on  $COV_i = COV$ ,  $f_i^{inc} \notin S_i$  and  $f_i^w = f$  (which is a reasonable assumption given that  $w_i$  can take any value on  $\mathbb{R}_+$ ), then we identify the constants  $\beta_f(COV) + \sum_{l=1}^L \gamma_l^{ws}(COV) \cdot CF_i^{ws}[l]$  for any given  $f$  and the parameters  $1/\lambda$  and  $\gamma_{l,f}^{ms}(COV)$  (for  $l = 1, \dots, L$  and  $f = 1, \dots, F$ ).<sup>24</sup> On the other hand, the constant  $\beta_{\tilde{f}}(COV) + \beta_{inc}$  and the parameters  $\gamma_l^{ws}(COV)$  ( $l = 1, \dots, L$ ) can be identified from the sub-sample with  $COV_i = COV$  and  $f_i^w = f_i^{inc} = \tilde{f}$  once we assume that the random vector  $(1, CF_i^{ws}[1], \dots, CF_i^{ws}[L])$  is of full rank conditional on  $COV_i = COV$  and  $f_i^w = f_i^{inc} = \tilde{f}$ . Once we have identified  $\gamma_l^{ws}(COV)$  ( $l = 1, \dots, L$ ), we also identify  $\beta_f(COV)$  for each  $f = 1, \dots, F$  and then also  $\beta_{inc}$ . We have thus identified all the coefficients in the augmented performance equation for  $COV \in \mathcal{S}_{COV}$ .

Now that the augmented performance equation is known, and given that  $\psi$  is identified from the bidding data, the incentive effect is also identified since  $\tau(w) = (\log(w/\psi(w)))/\lambda$ . Once  $\beta_f(COV)$  is known for each  $COV \in \mathcal{S}_{COV}$ , we identify  $\beta_x$ . Apart for the identification of this latter vector of parameters, we do not rely on variations in the covariates: The augmented performance equation (6) is identified for any given fixed set of covariates  $COV \in \mathcal{S}_{COV}$ . Also note that we are so far agnostic about how the covariates affect the distribution of signals and the  $\gamma$ -coefficients.

### 4.3 Two-stage parametric estimation

We present below the two-stage estimation procedure that is used in both our Monte Carlo study and empirical analysis. It relies on parametric assumptions regarding the incentive effect, the signal distributions, and how the payoff function depends on the observable covariates. More precisely, the following set of restrictions are imposed:

- The **linear-log specification** for the incentive effect:  $\tau(w) = \tau \cdot \log(w)$ , or equivalently  $\psi(w) = w^{1-\theta}$  with  $\theta = \lambda \cdot \tau < 1$ .
- **Gaussian signals**: the  $F + 1$  pairs of signal  $(s_{i,f}^{I,ms}, s_{i,f}^{P,ms})$ ,  $f = 1, \dots, F$  and  $(s_i^{I,ws}, s_i^{P,ws})$  are distributed independently according to bivariate normal distributions that do not depend on  $COV_i$ .
- **Symmetry across firms**: the distributions  $(s_{i,f}^{I,ms}, s_{i,f}^{P,ms})$ ,  $f = 1, \dots, F$  are identical across  $f$ .

<sup>24</sup>We use here that the control terms  $CF_i^{ws}[l]$  do not vary within the subsample with  $COV_i = COV$  and  $f_i^{inc} \notin S_i$ .

- **Linear payoff functions:**  $\overline{\mathcal{V}}(\mathbf{x}_{i,f}, \mathbf{z}_{i,f}, INC_{i,f}) = \alpha + \alpha_x \cdot \mathbf{x}_{i,f} + \alpha_z \cdot \mathbf{z}_{i,f} + \alpha_{inc} \cdot INC_{i,f}$ .

The multivariate normality assumption implies that both  $\lambda \cdot (s_{i,f}^{I,ms} + s_{i,f}^{P,ms})$  and  $\lambda \cdot (s_{i,f}^{I,ws} + s_{i,f}^{P,ws})$  are normally distributed. The distribution  $G_f^{ms}$  (resp.  $G^{ws}$ ) is thus a centered normal distribution and its variance is denoted  $\sigma_{ms}^2$  (resp.  $\sigma_{ws}^2$ ).<sup>25</sup> Our symmetric Gaussian structure also implies that  $\mathbb{E}[s_{i,f}^{P,ms} | s_{i,f}^{I,ms} + s_{i,f}^{P,ms} = x] = \gamma^{ms} \cdot x$  and  $\mathbb{E}[s_{i,f}^{P,ws} | s_{i,f}^{I,ws} + s_{i,f}^{P,ws} = x] = \gamma^{ws} \cdot x$ . Furthermore, if the correlation coefficient between  $s_{i,f}^{I,ms}$  and  $s_{i,f}^{P,ms}$  (resp.  $s_{i,f}^{I,ws}$  and  $s_{i,f}^{P,ws}$ ) is positive, then  $\gamma^{ms}$  (resp.  $\gamma^{ws}$ ) is necessarily positive and A7 is satisfied with  $L = 1$ .<sup>26</sup> The linear payoff function structure implies that

$$\overline{V}_{i,f} = \beta_f^* + \beta_x^* \cdot \mathbf{x}_{i,f} + \beta_z^* \cdot \mathbf{z}_{i,f} + \beta_{inc}^* \cdot INC_{i,f} \quad (7)$$

where  $\beta_f^* = \alpha + \beta_f + (\ln(E_{\epsilon_{i,f}}[e^{\lambda \cdot \epsilon_{i,f}}]))/\lambda$ ,  $\beta_x^* = \alpha_x + \beta_x$ ,  $\beta_z^* = \alpha_z$  and  $\beta_{inc}^* = \alpha_{inc} + \beta_{inc}$ . The primitives of interest of our auction model are thus fully determined by the (unconstrained) vector of parameters  $(\beta_f^*, \beta_x^*, \beta_z^*, \beta_{inc}^*)$ , the coefficient  $\theta < 1$  and the variances  $\sigma_{ms}^2$  and  $\sigma_{ws}^2$ . The first step of our estimation procedure consists in estimating these primitives through maximum likelihood, a procedure (detailed in the SA) which leads to the usual  $\sqrt{n}$ -asymptotic normality of the estimators.

To implement the second step we have to compute two control terms for each worker  $i$ ,  $CF_i^{ms}[1]$  and  $CF_i^{ws}[1]$ , shortly referred to as  $CF_i^{ms}$  and  $CF_i^{ws}$  from now on. Their precise forms depend on whether an incumbent is present among the potential auction participants, and, if there is an incumbent, on the identity of the winner. We start by detailing the computations for the case where the auction for worker  $i$  does not include an incumbent.

Let  $\Phi$  (resp.  $\phi$ ) denote the CDF (resp. PDF) of a standard normal distribution, and  $\widehat{s}_{i,f}(w)$  the cutoff aggregate signal that makes firm  $f$  bid exactly until  $w$ . This cutoff aggregate signal is thus the solution of  $w_{i,f}[NI](\widehat{s}_{i,f}(w)) = w$ , where the cutoff wage function  $w_{i,f}[NI] : \mathbb{R} \rightarrow \mathbb{R}_+$  is defined in (4). Under our parametric restrictions, the cutoff aggregate signal is thus characterized by

$$\lambda \cdot \widehat{s}_{i,f}(w) = (1 - \theta) \log(w) - \frac{\sigma_{ws}^2}{2} - \lambda \cdot \overline{V}_{i,f}. \quad (8)$$

If the auction outcome is  $(w_i, f_i^w)$ , we can infer that conditional on  $\mathcal{I}_i$ , the aggregate signal of the winning firm,  $\overline{s}_{i,f_i^w}$ , is distributed according to a centered normal distribution (with variance  $\sigma_{ms}^2$ ) truncated from below at  $\lambda \cdot \widehat{s}_{i,f_i^w}(w_i)$ . We obtain then:

$$CF_i^{ms} = \int_{\lambda \cdot \widehat{s}_{i,f_i^w}(w_i)}^{+\infty} u \cdot \frac{d\Phi\left(\frac{u}{\sigma_{ms}}\right)}{1 - \Phi\left(\frac{\lambda \cdot \widehat{s}_{i,f_i^w}(w_i)}{\sigma_{ms}}\right)} = \sigma_{ms} \cdot \frac{\phi\left(\frac{(1-\theta)\log(w) - \frac{\sigma_{ws}^2}{2} - \lambda \cdot \overline{V}_{i,f_i^w}}{\sigma_{ms}}\right)}{1 - \Phi\left(\frac{(1-\theta)\log(w) - \frac{\sigma_{ws}^2}{2} - \lambda \cdot \overline{V}_{i,f_i^w}}{\sigma_{ms}}\right)}. \quad (9)$$

<sup>25</sup>The variance  $\sigma_{ms}^2$  does not depend on  $f$  thanks to the symmetry restriction.

<sup>26</sup>The expression of the coefficient  $\gamma^{ms}$  as a function of the variances of  $s_{i,f}^{I,ms}$  and  $s_{i,f}^{P,ms}$  and their correlation coefficient, is given in the SA, and similarly for  $\gamma^{ws}$ .

The match-specific control function term corresponds to an inverse Mills ratio multiplied by the standard deviation  $\sigma_{ms}$ , and is similar to the correction term (to correct for sample selection bias) proposed originally by Heckman (1979). In auctions involving no incumbent, there is no updating on  $(s_i^{I,ws} + s_i^{P,ws})$ , and hence we have  $CF_i^{ws} = \int_{-\infty}^{+\infty} u \cdot d\Phi(u/\sigma_{ws}) = 0$ . From (9) we see that our control terms in the sub-sample of auctions without an incumbent depend on  $\lambda \cdot \bar{V}_{i,f_i^w}$ . This means that  $CF_i^{ms}$  varies in particular with  $z_{i,f_i^w}$ , contrary to the performance equation itself which does not depend on the auction shifters.

Let us now discuss the cases involving an incumbent whose computations are detailed in the SA. The subcase where the winner is the incumbent leads to control terms that are similar to the one in (9). On the contrary, the subcase where the winner is not the incumbent is quite different: on the one hand, the exact expressions of the control terms now depend on what is observed regarding the bidding history (beyond  $w_i$  and  $f_i^w$ ); on the other hand, the control terms do no longer solely depend on  $\bar{V}_{i,f_i^w}$ , implying that they vary with the covariates of non-winning bidders as well. For example, when the wage  $w_i^{inc}$  where the incumbent has dropped-out is publicly observed, we can infer that its aggregate signal is equal to  $(1 - \theta)\log(w_i^{inc}) - \lambda \cdot \bar{V}_{i,f_i^{inc}}$ , which further implies the following expression of the worker-specific control term:

$$CF_i^{ws} = \frac{\sigma_{ws}^2}{\sigma_{ms}^2 + \sigma_{ws}^2} \cdot [(1 - \theta)\log(w_i^{inc}) - \lambda \cdot \bar{V}_{i,f_i^{inc}}]. \quad (10)$$

As detailed in the SA, the corresponding match-specific control term depends on both  $\bar{V}_{i,f_i^w}$  and  $\bar{V}_{i,f_i^{inc}}$ .

The second step of our estimation procedure consists in estimating, by OLS, the performance equation augmented with the estimated control terms:

$$y_{i,f_i^w} = \beta_{f_i^w} + \beta_x \cdot \mathbf{x}_{i,f_i^w} + \beta_{inc} \cdot INC_{i,f_i^w} + \tau \cdot \log(w_i) + \gamma^{ms} \cdot \widehat{CF}_i^{ms} + \gamma^{ws} \cdot \widehat{CF}_i^{ws} + \xi_{i,f_i^w}. \quad (11)$$

Here  $\widehat{CF}_i^{ms}$  and  $\widehat{CF}_i^{ws}$  are the estimated control functions, i.e., the expressions obtained after replacing all unknown parameters in the control functions by their first-stage estimates. The resulting estimator is consistent and asymptotically normally distributed (see Wooldridge (2002)). The standard errors are obtained using a percentile bootstrap method.<sup>27</sup> Combining our first-stage estimates of the  $\beta^*$ -coefficients with our second-stage estimates of the  $\beta$ -coefficients, leads to estimates of  $\alpha_x$ ,  $\alpha_z^*$  and  $\alpha_{inc}$ .

From the exact expression of the control terms in this parametric setting, we see that these terms depend non-linearly on the regressors appearing in the performance equation (1), but also on the covariates  $z_{i,f_i^w}$ , and  $(\mathbf{x}_{i,f_i^{inc}}, z_{i,f_i^{inc}})$  (when  $f_i^{inc} \neq f_i^w$  and  $f_i^{inc} \in S_i$ ). Their presence constitutes a source of variation in the control function terms which in practice reduces

<sup>27</sup>Alternatively, the standard errors could be obtained using the two-stage asymptotic variance expression (see Wooldridge (2002)). The advantage of the percentile bootstrap method, when the sample size is small, is that it does not assume normality of the estimated parameter distribution. Another advantage is that it does not require a variance estimator.

collinearity among the explanatory variables in the augmented performance equation, thereby augmenting the precision of the estimates.

#### 4.4 Our two-stage methodology under private values

Let us now briefly discuss how our methodology would change when the auction data never involve an incumbent. Proposition G.1 in the SA establishes that the primitives of the bidding model are no longer identified from the auction data: in particular, they can always be rationalized by a private value model without any incentive effect (i.e., with  $s_i^{I,ws} = s_i^{P,ws} = 0$  and  $\tau(\cdot) \equiv 0$ ). A possible avenue for identification is to impose a parametric restriction on the incentive effect: if we assume the linear-log specification  $\tau(w) = \tau \cdot \log(w)$  as in Section 4.3, then we can develop a similar two-stage identification strategy as above. We can identify the distribution of  $\lambda \cdot \bar{s}_{i,f_i^w} / (1 - \theta)$  conditional on  $\mathcal{I}_i$  and, as argued in Section 4.2, plug the corresponding moments of this conditional distribution as match-specific control terms in the performance equation. We then identify  $\beta_x$  and  $\tau$  from the following augmented performance equation:<sup>28</sup>

$$y_{i,f_i^w} = \beta_{f_i^w} + \beta_x \cdot x_{i,f_i^w} + \tau \cdot \log(w_i) + \sum_{l=1}^L \gamma_{l,f_i^w}^{ms} (COV_i) \cdot CF_i^{ms}[l] + \xi_{i,f_i^w}. \quad (12)$$

In our Monte Carlo simulations, this approach is referred to as the PV methodology. Note that, if in addition  $\tau = 0$ , then the PV methodology is highly related to the polychotomous selection problem addressed by Dubin and McFadden (1984) through control functions.

The PV methodology leads to consistent estimates if the auctions never involve an incumbent. It also produces consistent estimates if the incumbent does participate but never observes the worker-specific signals (or equivalently if  $s_i^{I,ws} = s_i^{P,ws} = 0$ ). However, there is no guarantee that our general methodology (referred to next as the CV methodology) remains consistent in these settings. The CV method may perform poorly because the control terms  $CF_i^k[l]$ ,  $k = ms, ws$  then fail to be identified from the bidding data. Whether it is better to use the PV or CV methodology is an empirical question addressed in Section 5.

## 5 Simulation Study

The aim of the Monte Carlo exercise is to explore how our methodology performs in finite samples. The simulation is designed to broadly match the setting of our empirical application. There are eight firms ( $F = 8$ ), the firm-worker characteristics are assumed to be uni-dimensional and the same across all firms  $f$ ,  $x_{i,f} \equiv x_i \in \mathbb{R}$ , and the auction shifters are also uni-dimensional but vary across firms,  $z_{i,f} \in \mathbb{R}$ . Half of each simulated sample is composed of

<sup>28</sup>Compared to (6), the term  $(\log(w_i/\psi(w_i))/\lambda)$  has been replaced by  $\tau \cdot \log(w_i)$ , and, naturally, the variable  $INC_{i,f_i^w}$  and the worker-specific control function term  $CF_i^{ms}[l]$  are absent. Note that the parameter  $\lambda$  is no longer identified from the second stage.

auctions without incumbents while the other half is made up of auctions with an incumbent (picked randomly). The main features of our calibration choices are as follows (some of the details are provided in the SA). We assume that the incentive effect takes the form  $\tau(w) = \tau \log w$ , where  $\tau$  takes the values  $-0.8, -0.5, 0, 0.5$  and  $0.8$ . The variables  $x_i$  and  $z_{i,f}$ ,  $f = 1, \dots, F$  and the noise  $\epsilon_{i,f}$  are assumed to be independently distributed according to a centered normal distribution with variance equal to 1. We set  $\beta_f = 0$  for all bidders,  $\beta_{inc} = 0$  and  $\beta_x = 1$ . We assume that the intrinsic signals  $s_{i,f}^{I,ms}$  and  $s_i^{I,ws}$  are equal to zero while the productivity signals  $s_{i,f}^{P,ms}$  and  $s_i^{P,ws}$  follow centered normal distributions with variances  $\sigma_{ms}$  and  $\sigma_{ws}$ , respectively. For all simulations we set  $\sigma_{ms} = 1 - \tau$  (resp.  $\sigma_{ws} = (1 - \tau) \cdot \tilde{\sigma}$ ), while  $\tilde{\sigma}$  takes the values 0, 1 and 2. The parameter  $\tilde{\sigma} = \sigma_{ws}/\sigma_{ms}$  captures the relative strength of the common value signal compared to the private value signal in bidders' valuations ( $\tilde{\sigma} = 0$  corresponds to a pure private value environment as discussed in Section 4.4).

Table 1A (resp. 1B) reports the mean estimates of the parameter  $\tau$  (resp.  $\beta_x$ ) over 1,000 Monte Carlo replications of the parameters, using a sample size of  $N=300$ ,<sup>29</sup> for three methodologies: 1) standard OLS; 2) the PV methodology described in Section 4.4; and 3) our general methodology incorporating both private and common values and which was described in Section 4.3 (CV methodology). The PV methodology amounts to treating the incumbents in the simulated samples as if they are non-incumbents, i.e., we apply the estimation strategy outlined in Section 4.4 as if all firms are non-incumbents. For both the PV and CV methodology we fix  $L = 1$  which is consistent with the Gaussian structure. We also report the percentage of the players that are sold: this percentage varies between 0.28 and 0.98 in our simulations. Table 1C (resp. 1D) reports the estimated lower and upper bounds of the 95% confidence intervals for  $\tau$  (resp.  $\beta_x$ ) for the CV methodology. We estimate these bounds using the Warp-Speed method developed by Giacomini et al. (2013) and briefly summarized in the SA.<sup>30</sup> Also reported in Table 1C and 1D are estimates of statistical power and the empirical coverage probabilities for the parameters  $\tau$  and  $\beta_x$ .

We see from Table 1A that, as predicted in Section 2, the estimator of  $\tau$  is upward biased. From Table 1B we see that the bias is also very substantial for the parameter  $\beta_x$ . For instance, when  $\tau = -0.8$ , the mean OLS estimate of  $\tau$  ranges between  $-0.23$  (when  $\tilde{\sigma} = 0$ ) and  $0.16$  (when  $\tilde{\sigma} = 1$ ), while the mean OLS estimate of  $\beta_x$  (recall that its true value is 1) varies between  $0.10$  ( $\tilde{\sigma} = 1$ ) and  $0.42$  ( $\tilde{\sigma} = 0$ ). We also see that the smallest biases are always obtained when  $\tilde{\sigma} = 0$ . For  $\tau \geq 0$ , the biases increase in  $\tilde{\sigma}$ , while for the cases  $\tau = -0.5$  and  $-0.8$  the strongest biases are obtained for the intermediary value  $\tilde{\sigma} = 1$ . It is notable that the biases under OLS for both  $\tau$  and  $\beta_x$  are strong when  $\tau = 0$ , even if  $\tilde{\sigma} = 0$ . They are comparable in magnitude to the biases obtained with other values for  $\tau$ : in the absence of wage endogeneity the bias due to non-random matching and selection remains strong, even if the incumbency status plays no role in the data generation process. The PV estimator performs well when  $\tilde{\sigma} = 0$ , i.e., when the data are generated with pure private values. By contrast, the results are biased when  $\tilde{\sigma} > 0$

<sup>29</sup>Additional results with  $N=1,000$  are reported in Table K.1 in the SA.

<sup>30</sup>By relying on a single bootstrap sample for each replication, the Warp-Speed method reduces drastically the computational cost of our simulations.

Table 1: Monte Carlo Results for  $N = 300$

| Table 1A: Mean of $\hat{\tau}$          |                  |       |       |       |       |       |       |       |       |       |      |      |      |      |      |      |
|---|------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|------|------|------|------|------|------|
|   |                  | -0.8  |       |       | -0.5  |       |       | 0     |       |       | 0.5  |      |      | 0.8  |      |      |
| $\tau =$                                | $\hat{\sigma} =$ | 0     | 1     | 2     | 0     | 1     | 2     | 0     | 1     | 2     | 0    | 1    | 2    | 0    | 1    | 2    |
| OLS                                     |                  | -0.22 | 0.16  | -0.12 | -0.02 | 0.33  | 0.19  | 0.31  | 0.59  | 0.65  | 0.64 | 0.82 | 0.96 | 0.85 | 0.95 | 1.02 |
| PV                                      |                  | -0.78 | -0.45 | -0.41 | -0.48 | -0.2  | -0.09 | 0.01  | 0.19  | 0.36  | 0.49 | 0.6  | 0.73 | 0.78 | 0.85 | 0.9  |
| CV                                      |                  | -0.82 | -0.81 | -0.93 | -0.53 | -0.5  | -0.62 | -0.02 | 0.01  | -0.08 | 0.48 | 0.49 | 0.45 | 0.79 | 0.78 | 0.75 |
|   |                  |       |       |       |       |       |       |       |       |       |      |      |      |      |      |      |
| Table 1B: Mean of $\hat{\beta}_x$       |                  |       |       |       |       |       |       |       |       |       |      |      |      |      |      |      |
|   |                  | -0.8  |       |       | -0.5  |       |       | 0     |       |       | 0.5  |      |      | 0.8  |      |      |
| $\tau =$                                | $\hat{\sigma} =$ | 0     | 1     | 2     | 0     | 1     | 2     | 0     | 1     | 2     | 0    | 1    | 2    | 0    | 1    | 2    |
| OLS                                     |                  | 0.42  | 0.1   | 0.4   | 0.52  | 0.23  | 0.4   | 0.68  | 0.46  | 0.43  | 0.84 | 0.72 | 0.62 | 0.92 | 0.88 | 0.83 |
| PV                                      |                  | 0.95  | 0.68  | 0.63  | 0.97  | 0.74  | 0.61  | 0.99  | 0.85  | 0.68  | 1    | 0.93 | 0.82 | 1.01 | 0.98 | 0.93 |
| CV                                      |                  | 1.05  | 1.08  | 1.19  | 1.04  | 1.06  | 1.17  | 1.03  | 1.02  | 1.1   | 1.01 | 1.01 | 1.04 | 0.99 | 1.01 | 1.02 |
|   |                  |       |       |       |       |       |       |       |       |       |      |      |      |      |      |      |
| Table 1C: Power and CI Bounds $\tau$    |                  |       |       |       |       |       |       |       |       |       |      |      |      |      |      |      |
|   |                  | -0.8  |       |       | -0.5  |       |       | 0     |       |       | 0.5  |      |      | 0.8  |      |      |
| $\tau =$                                | $\hat{\sigma} =$ | 0     | 1     | 2     | 0     | 1     | 2     | 0     | 1     | 2     | 0    | 1    | 2    | 0    | 1    | 2    |
| Power                                   |                  | 1.00  | 0.94  | 1.00  | 0.96  | 0.74  | 0.95  | 0.05  | 0.05  | 0.08  | 0.91 | 0.8  | 0.73 | 0.86 | 0.69 | 0.94 |
| Lower Bound                             |                  | -1.14 | -1.25 | -1.28 | -0.82 | -0.91 | -0.98 | -0.38 | -0.37 | -0.43 | 0.19 | 0.15 | 0.11 | 0.26 | 0.15 | 0.32 |
| Upper Bound                             |                  | -0.54 | -0.35 | -0.59 | -0.25 | -0.13 | -0.28 | 0.23  | 0.34  | 0.25  | 0.78 | 0.84 | 0.76 | 1.32 | 1.37 | 1.21 |
| Empirical Coverage                      |                  | 0.94  | 0.96  | 0.9   | 0.94  | 0.94  | 0.9   | 0.95  | 0.95  | 0.92  | 0.96 | 0.94 | 0.94 | 0.96 | 0.95 | 0.96 |
|   |                  |       |       |       |       |       |       |       |       |       |      |      |      |      |      |      |
| Table 1D: Power and CI Bounds $\beta_x$ |                  |       |       |       |       |       |       |       |       |       |      |      |      |      |      |      |
|   |                  | -0.8  |       |       | -0.5  |       |       | 0     |       |       | 0.5  |      |      | 0.8  |      |      |
| $\tau =$                                | $\hat{\sigma} =$ | 0     | 1     | 2     | 0     | 1     | 2     | 0     | 1     | 2     | 0    | 1    | 2    | 0    | 1    | 2    |
| 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  | 0.66  | 0.75  | 0.62  | 0.72  | 0.77  | 0.69  | 0.67  | 0.76 | 0.71 | 0.70 | 0.64 | 0.67 | 0.68 |
| Upper Bound                             |                  | 1.39  | 1.59  | 1.77  | 1.38  | 1.53  | 1.68  | 1.30  | 1.41  | 1.47  | 1.28 | 1.30 | 1.39 | 1.32 | 1.38 | 1.40 |
| Empirical Coverage                      |                  | 0.92  | 0.95  | 0.90  | 0.94  | 0.94  | 0.88  | 0.94  | 0.94  | 0.93  | 0.95 | 0.95 | 0.95 | 0.96 | 0.94 | 0.96 |
|   |                  |       |       |       |       |       |       |       |       |       |      |      |      |      |      |      |
| Proportion Sold                         |                  | 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 |



but less so compared to OLS. The PV method reduces roughly by half (resp. two thirds) the bias of the OLS estimator when  $\tilde{\sigma} = 2$  (resp.  $\tilde{\sigma} = 1$ ). Finally, we find that the performance of the CV estimator is comparable to that of the PV estimator when  $\tilde{\sigma} = 0$ , which is surprising as there is no formal guarantee that our CV methodology is consistent when the true model is a pure private value one, in particular because the function  $\psi$  (or equivalently the parameter  $\theta$  in our linear-log parametric specification) is no longer identified. However, the CV estimator clearly outperforms the PV estimator in the case  $\tilde{\sigma} > 0$ . When  $\tilde{\sigma} = 1$ , the CV estimator of  $\tau$  is perfectly consistent for all simulations; the parameter  $\beta_x$  is also well estimated albeit with a small upward bias. When  $\tilde{\sigma} = 2$ , the CV estimator of  $\tau$  (resp.  $\beta_x$ ) is slightly downward (resp. upward) biased, but the biases are much smaller than for the PV estimator.

Tables 1C and 1D show 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. Our tests of the null hypothesis  $H_0 : \tau = 0$  have also high power. The size is adequate when  $\tilde{\sigma}$  equals 0 or 1, but seems a bit too large (0.8) when  $\tilde{\sigma} = 2$ . 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  $\sigma_{ws}$ . The estimated lower-bounds and upper-bounds of the CIs are symmetrically distributed around the estimated mean values. We also observe that the CIs for  $\beta_x$  become tighter as  $\tilde{\sigma}$  gets smaller. In contrast, there is not a clear-cut relationship between the CIs for  $\tau$  and  $\tilde{\sigma}$ : for  $\tau = -0.8, -0.5, 0$ , the CI is tighter as  $\tilde{\sigma}$  gets smaller, while the reverse holds for  $\tau = 0.5, 0.8$ . Intuitively, there are two counter-vailing forces when  $\tilde{\sigma}$ , or equivalently  $\sigma_{ws}$  for a given  $\tau$ , gets larger: on the one hand, the selection effect becomes more important (as illustrated by the stronger bias of OLS estimates) and introduces thus additional noise in the performance equation reducing the precision of the estimates. On the other hand,  $\widehat{CF}_i^{ws}$  becomes a less noisy estimator of the control term  $CF_i^{ws}$  (in particular because common values helps the identification of  $\psi$ ) thereby increasing precision. Depending on the parameter being estimated one of these forces dominates the other.

## 6 Data and Empirical Application

### 6.1 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. We obtained the performance data

on all matches played in the tournament and the auction data from [www.espnricinfo.com](http://www.espnricinfo.com).<sup>31</sup>

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. One feature of cricket, as in a few other sports such as baseball, is that 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. In order to derive a performance measure for each individual player from various batting and bowling statistics observed individually during the tournament, we use an existing methodology which ranks players in the IPL.<sup>32</sup> The way this performance index is built is described in the SA where we also give additional information on the rules of the cricket game. The main steps are as follows: 1) points are awarded to players separately for batting and bowling across every game; 2) these points are then accumulated over the tournament and multiplied by a speciality specific efficiency factor (the relative strike rate for batting and the relative economy rate for bowling, respectively); 3) the resulting batting and bowling performance measures are then normalized by the number of games played by each player.<sup>33</sup>

## 6.2 Player Auctions and Wage Contract

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.

**Player Contracts:** The gross wage of a player is equal to his auction price.<sup>34</sup> In the 2014 auction, player wage contracts were fixed for a one-year term with the option of renewal for an additional one or two years. Players whose contracts were terminated at the end of the first year would be pooled into a mini-auction in the subsequent year and re-allocated across teams through this auction. As a result, the auctions and tournament present an ideal setting for our analysis as we are able to examine player performance in the first season following the auction, where players face a genuine incentive to perform that is effective immediately after their wage-determination.

**Auction format:** 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. 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

---

<sup>31</sup>All data on player auctions were manually compiled from the recordings and minutes of the (publicly broadcast) auction proceedings.

<sup>32</sup>See <https://bit.ly/2CvCB44> for a description.

<sup>33</sup>The performance measure for all-rounders (who both bat and bowl) is the sum of their batting and bowling performance measure.

<sup>34</sup>All player salaries are taxed in India; however overseas players face a (uniformly) lower tax burden on their salary compared to Indian players (approximately 10% compared to 40%) In the analysis, our measure of wages are in gross terms, but we account for the tax-differential by including a dummy for Indian players.

among them being the player’s past performance.<sup>35</sup> 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).<sup>36</sup> 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 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.<sup>37</sup> Next, to simplify, we do not put much emphasis on the RTM option rule (which plays a role in about one fourth of our auction sample) but we do take it into account in a structural way. See the SA for details.

The auctioneer arranges the players into different ‘sets’ by their cricketing speciality, quality, 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. In our empirical specification, we explicitly control for the set in which each player appears (using a dummy variable for the player’s set) which absorbs any player-quality specific effect that may be related to the predetermined set ordering. This leaves, therefore, only the exogenous variation from the within-set player ordering which is uncorrelated with player quality given that it is obtained from a (random) lottery.

**Auction rules:** 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. These rules include: 1) A spending cap: 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) An overseas player quota: 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.

**Auction shifter variables:** Our analysis deals with the sequential aspect of the game through a reduced form approach by including the auction shifters  $z_{i,f}$  into the specification

---

<sup>35</sup>In 2014, seven different reserve prices were used from 1 up to 20 Millions of Rupees.

<sup>36</sup>Lamy et al. (2016) do take into account increments in their auction model but at the cost of being not fully structural.

<sup>37</sup>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 which has the possibility to exert the RTM option was allowed to bid in the auction exactly as the other bidders.

of  $V_{i,f}$ . This vector can include e.g. the amount of money spent by  $f$  in the auctions preceding the auction for agent  $i$ , the number and characteristics of agents bought prior to  $i$  (capturing possible substitutabilities or complementarities between different types of agents), the order in the auction sequence and so forth. A priori there is no reason why these shifters would directly affect player performance. For instance, the characteristics of previously purchased agents is not expected to play a role given the individualized nature of cricket. Similarly, order of sale (and hence the remaining budget which is a function of order) should not matter given that is randomly drawn within each set.

### 6.3 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.<sup>38</sup> 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 the SA.<sup>39</sup> For each auction we 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). Table 2 present summary statistics from the perspective of both teams and players.

The upper panel of Table 2 shows that approximately half of all auctions (156 out of 317) involved an incumbent. The probability that the player is sold in an auction with (resp. without) an incumbent equal to 58% (resp. 19%). Among auctions with an incumbent, the probability that the incumbent (resp. a non-incumbent) becomes the winner is equal to 24% (5%). As indicated in the first 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. Given the number of players retained by teams 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), teams were allocated an average budget of 5.65 million USD for purchasing players. On average, bidders consume 90% of their allocated budget.

The two lower panels of Table 2 contain summary statistics on the auction data, first for the full sample, and then for the players who were sold and performed on the field. For this latter sample, we also reports statistics on our composite measure of performance and the wage

---

<sup>38</sup>A newcomer is a player who has never been called by his national team. Such players may have experience playing in past IPL tournaments but we nonetheless use the terminology "Newcomer" as this is the official designation for such players.

<sup>39</sup>For 17 sold players we do not observe the performance measure because they do not perform on the field.

Table 2: Summary statistics

| Variable   | Description  | # Auctions | Percent of total |
|--|--|------------|------------------|
| <i>Bidder incumbency status:</i>                         |  |            |                  |
| # Auctions without incumbent                             | Number of auctions without incumbent or with an incumbent who is not eligible to bid   | 161        | 50.79            |
| # 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            |
| <i>Bidder purchases and constraints in the auctions:</i> |  |            |                  |
| # 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             |
| 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             |
| <i>Full sample descriptives:</i>                         |  |            |                  |
| # active bidders   | Participating bidders for each player auction  | 0.87       | 1.29             |
| Reserve price  | Reservation wage set by auctioneer   | 0.10       | 0.09             |
| Order of sale  | 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             |
| <i>Performance sample descriptives:</i>                  |  |            |                  |
| # 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 of sale  | 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., $\sum_j INC_j$ )  | 0.74       | 0.43             |
| Winner is incumbent                                      | Dummy indicating whether player was matched with the winner in the previous season (formally it corresponds to $INC_{j^w}$ ) | 0.30       | 0.46             |
| Tournament Points <sup>†</sup>                           | 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).

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.

#### 6.4 Empirical Specification and Estimation Results

We consider the parametric restrictions from Section 4.3. 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,<sup>40</sup> 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  $INC_{i,f}$  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  $\sum_f INC_{i,f}$ ). The vector of auction shifters  $z_{i,f}$  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.

---

<sup>40</sup>Included are dummies for batsman and bowler. There are too few wicket-keepers in the sold-player sample to add yet another speciality dummy.

Table 3: Empirical results

|   | OLS                     | First Stage                | Second Stage                |
|---|-------------------------|----------------------------|-----------------------------|
| Log Wage ( $\tau$ )                                   | 5.80***<br>[2.50, 9.07] |                            | 4.14**<br>[0.15, 8.11]      |
| <b>Team-player characteristics (<math>x</math>):</b>  |                         |                            |                             |
| Indian  | -4.68<br>[-10.66, 1.30] | -2.52***<br>[-3.35, -1.68] | -3.22<br>[-10.31, 4.48]     |
| Newcomer  | -1.17<br>[-9.54, 7.20]  | 0.51<br>[-0.10, 1.13]      | -2.43<br>[-11.58, 5.71]     |
| Speciality: Batsman                                   | 2.16<br>[-5.76, 10.08]  | 2.16***<br>[1.12, 3.20]    | 3.82<br>[-3.02, 11.88]      |
| Speciality: Bowler                                    | 8.52**<br>[1.64, 15.39] | 3.01***<br>[2.01, 4.02]    | 10.15***<br>[3.47, 17.97]   |
| Incumbent present                                     | 1.87<br>[-4.57, 8.31]   | 2.57***<br>[1.27, 3.87]    | 5.48<br>[-3.05, 15.49]      |
| Bidder is incumbent                                   |                         | 2.23***<br>[1.23, 3.21]    |                             |
| Winner is incumbent                                   | -2.27<br>[-8.26, 3.71]  |                            | -15.27**<br>[-30.80, -1.47] |
| <b>Auction shifters (<math>z</math>):</b>             |                         |                            |                             |
| Incumbent present & no RTM card                       |                         | -1.33***<br>[-2.01, -0.65] |                             |
| Order of sale   |                         | -0.01<br>[-0.06, 0.05]     |                             |
| Remaining budget (in logs)                            |                         | 0.55***<br>[0.29, 0.82]    |                             |
| # Batsman bought                                      |                         | -0.36***<br>[-0.58, -1.14] |                             |
| # Bowlers bought                                      |                         | -0.48***<br>[-0.64, -0.32] |                             |
| # Wicket-keepers bought                               |                         | 0.01<br>[-0.41, 0.42]      |                             |
| # All-rounders bought                                 |                         | 0.41***<br>[0.24, 0.58]    |                             |
| # Overseas players bought                             |                         | -0.63***<br>[-0.80, -0.46] |                             |
| <b>Other Structural Parameters:</b>                   |                         |                            |                             |
| $\gamma^{ms}$   |                         |                            | -0.47<br>[-3.82, 2.05]      |
| $\gamma^{ws}$   |                         |                            | 4.01**<br>[0.30, 16.80]     |
| $\theta$  |                         | 0.77***<br>[0.60, 0.94]    |                             |
| $\sigma_{ms}^* \equiv \frac{\sigma_{ms}}{(1-\theta)}$ |                         | 1.48***<br>[0.75, 2.92]    |                             |
| $\sigma_{ws}^* \equiv \frac{\sigma_{ws}}{(1-\theta)}$ |                         | 2.22***<br>[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  $\beta_x^*$ ,  $\beta_{inc}^*$ ,  $\beta_z^*$ ,  $\theta$ ,  $\sigma_{ms}$ , and  $\sigma_{ws}$ , and 95% CIs based on the asymptotic ML standard errors. Column 3 reports OLS estimates of the parameters in the augmented performance equation (11), 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  $\gamma^{ms}$  and  $\gamma^{ws}$  where it is one-sided given the restriction A9 and 90% CIs are reported in parentheses.

Table 3 contains our empirical results. The first column gives OLS estimates of the parameters appearing in the uncorrected performance equation (1). The results indicate that a player's wage has a positive and statistically significant 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, except for the player speciality bowler, the indicators for incumbency status, nationality, newcomers, and batsmen are not significant.

The next two columns give the estimation results corresponding to our methodology. Column 2 contains the first stage results, i.e., the ML estimates of the auction primitives  $\beta_{inc}^*$ ,  $\beta_x^*$ ,  $\beta_z^*$ ,  $\theta$ ,  $\sigma_{ms}/(1-\theta)$ , and  $\sigma_{ws}/(1-\theta)$ , and the 95% CIs based on the asymptotic ML standard errors. For the purpose of the discussion below, we recall that team  $f$ 's valuation for player  $i$  corresponds to  $w_{i,f}^*$ , that is to say the amount that  $f$  is prepared to bid if it could observe the worker-specific signals  $s_i^{P,ws}$  and  $s_i^{I,ws}$ .

The lower panel of the table shows that the first stage estimate of  $\theta$  is 0.77 (significant at the 1% level). Furthermore, the estimates of  $\sigma_{ms}/(1-\theta)$  and  $\sigma_{ws}/(1-\theta)$  are 1.48 and 2.22. From the estimated  $\beta^*$ -coefficients, our estimate for the standard deviation of  $\lambda\bar{V}_{i,f}/(1-\theta)$  is 2.2. Since  $\log(w_{i,f}^*) = \lambda[\bar{V}_{i,f} + s_{i,f}^{P,ms} + s_{i,f}^{I,ms} + s_i^{P,ws} + s_i^{I,ws}]/(1-\theta)$ , and using the independence between the covariates (appearing in  $\bar{V}_{i,f}$ ), and the match-specific and worker-specific signals, our estimate for the standard deviation of  $\log(w_{i,f}^*)$  is equal to 3.4. On the whole, we obtain that the variance of  $\lambda\bar{V}_{i,f}/(1-\theta)$  explains about 40% of the variance of  $\log(w_{i,f}^*)$ , indicating that private signals matter a lot in determining the variation in valuations. We also obtain that a standard deviation increase in the aggregate signal of a non-incumbent (resp. the incumbent) increase the logarithm of its valuation by an amount which represents 0.44 (resp. 0.79) of its standard deviation.

All team-player characteristics are statistically significant except the newcomer indicator: teams reduce their valuation for Indian players and increase it for certain player specialties (batsmen, bowlers); our results also indicate that 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).<sup>41</sup> Specifically, a player who participated in the tournament previously significantly increases the associated teams' valuation by 0.75 standard deviations relative to other players when his incumbent team has still a RTM card.<sup>42</sup> 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 valuation and the valuation of a non-incumbent corresponds to 0.4 standard deviations, which is consistent with the observation that the incumbent's winning probability is

<sup>41</sup>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 valuations but also through the asymmetric information across bidders w.r.t. commonly valued signals and possibly through the RTM option.

<sup>42</sup>Given the estimated coefficient on the auction shifter 'Incumbent present & no RTM card', the corresponding increase when the incumbent team has exhausted her RTM cards is only 0.37 standard deviations.



five times larger than for a given non-incumbent.

Let us next look at the results concerning the auction shifters  $z$ . The order of sale within the set does not significantly affect teams' 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 valuation for the given player by 0.015 standard deviations. The five last auction shifters 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 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 (11), 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 5% 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 upward 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 either efficiency wage (Shapiro and Stiglitz (1984)) or fairness/reciprocity (Akerlof and Yellen (1990)) theories.

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  $\beta_x^*$  and  $\beta_x$  are (significantly) of different signs: 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  $\alpha_{inc}$  in terms of parameters that are identified from our first and second stages, our estimate for  $\alpha_{inc}$  is 3.4 and is significant 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  $CF^{ws}$  is statistically significant (at the 5% level with a one-sided test), but the one associated with  $CF^{ms}$  is not. The fact that (at least) one of our control functions is significant confirms that endogeneity is an issue in our application.

As a robustness check, we also report results from specifications that include a player's past IPL performance in the set of player characteristics when this information is available.<sup>43</sup> It allows to control for player-specific signals commonly observed by all bidders that would challenge the independence assumption made in A1 ii) and on which our methodology is relying. The results, reported in the SA Table M.2, are robust to the inclusion of such player's past performance and both the magnitude and direction of the incentive and selection effects remain unchanged. It should be noted that, as discussed previously in section 6, that the *newcomer* and set-specific fixed effects partly already account for a dimension of the player's past performance as players are organized into sets depending on their quality.

## 6.5 Counterfactuals

In this section we use the estimated parameters from the first-stage of our empirical exercise to evaluate (i) the efficiency of matching between players and teams, (ii) the effect of changing reservation wages on the probability that players are sold and their expected revenues, and (iii) the effect of excluding the incumbent on the sale probability and expected revenues. In all simulations presented below, we put aside the sub-sample of auctions wherein there is an incumbent who benefits from an RTM option (but keep those wherein there is an incumbent who does not benefit from this option).

### 6.5.1 Matching Efficiency under Asymmetric Information

We designate a team-player match as efficient if player  $i$  is matched to the team with the highest valuation whenever it exceeds the reservation wage, or if  $i$  remains unsold whenever it is below. Formally, following the notation introduced in Section 3, it is efficient to match team  $f' \in \{1, \dots, F\}$  with player  $i$  if  $f' \in \text{Arg max}_{f \in \{1, \dots, F\}} w_{i,f}^*$  and  $w_{i,f'}^* \geq W_i^r$ , and to leave player  $i$  unmatched if  $\max_{f \in \{1, \dots, F\}} w_{i,f}^* < W_i^r$ . Note that our equilibrium analysis guarantees that the winning bidder (if any) always belongs to  $\text{Arg max}_{f \in \{1, \dots, F\}} w_{i,f}^*$  in auctions without incumbents. In contrast, if there is an incumbent, this firm may win the auction although it does not have the highest valuation and, conversely, a non-incumbent may win the auction although the incumbent has a higher valuation.<sup>44</sup> Auctions with an incumbent thus involve an additional potential source of inefficiency that does not exist in auctions without an incumbent.

We estimate our (in)efficiency measure using simulations. For each  $i$ , we take the variables  $(x_{i,f}, z_{i,f}, INC_{i,f})$  for all  $f$ , and  $W_i^r$ , as actually observed in the data, and simulate the vector

<sup>43</sup>More precisely, we add two variables to the specification: (i) a dummy variable taking the value one if past IPL performance is available for the player, and (ii) the interaction of this dummy variable and the previous three years' average for past IPL performance.

<sup>44</sup>Nevertheless, when a non-incumbent wins, it has necessarily the highest valuation among all non-incumbents.

Table 4: Percentage of inefficient matching

| Sample:                            | Inefficient | Sources of Inefficiency |      |      |      |
|------------------------------------|-------------|-------------------------|------|------|------|
|                                    |             | (A)                     | (B)  | (C)  | (D)  |
| Without Incumbent                  | 0.27        | 0.11                    | –    | –    | 0.16 |
| With Incumbent (and no RTM option) | 0.22        | 0.04                    | 0.04 | 0.04 | 0.09 |

Note: Sources of inefficiency are as follows - (A) sale whereas no-sale is efficient; (B) sale to incumbent whereas sale to non-incumbent is efficient; (C) sale to non-incumbent whereas sale to incumbent is efficient; (D) no-sale whereas sale is efficient.

of private signals  $\{s_{i,f}\}_{f=1,\dots,F}$  10,000 times. For each simulation, we compute the auction outcome and teams' valuations which enable us to determine whether the simulated match is efficient or inefficient. We then calculate the fraction of times (across 10,000 simulations) that the matching is inefficient for  $i$ . Table 4 gives then the average of these estimated probabilities that the final assignment fails to be efficient for the sub-samples of auctions with and without incumbents and details the sources of inefficiencies.

We observe that matching inefficiencies are not negligible: In the sample without incumbents, inefficiencies arise in 27% of our simulations. In the sample with an incumbent, we find that the inefficiency probability is lower (22%). The fact that the latter type of auction is relatively efficient results from two opposing effects. On the one hand, as remarked above, auctions with an incumbent introduce an additional form of inefficiency. Indeed, columns B and C in Table 4 indicate that for 8% of the outcomes in the sample with an incumbent, the winning team is not the bidder having the highest valuation. On the other hand, the presence of an incumbent allows to reduce the probability that a player remains unassigned although it would have been efficient to assign him to a team (column D),<sup>45</sup> and, conversely, the probability that the player is assigned to a team although it would have been efficient to keep him unassigned (column A). These reductions arise thanks to the informational linkage conveyed by Bayesian updating from observing the incumbent's bidding behavior. Here the positive effect on efficiency due to the informational linkage thus dominates the negative effect due to the inadequate selection of auction winners.

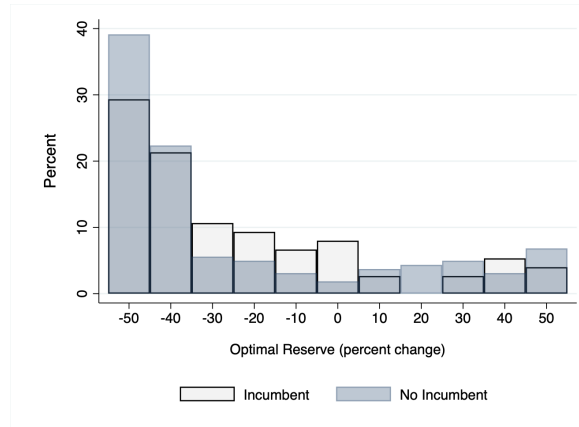
Note that our results strongly contrast with [Carnehl and Weiergraeber \(2022\)](#)'s structural analysis of German procurement auctions for railway services. Their auctions for the so-called net contracts (i.e., for which ticket revenues accrue to the contractor) involve informational asymmetries (similarly as in our auctions with an incumbent), and they find a much higher inefficiency rate than we do (their estimates exceed 80%). An explanation for this discrepancy is that their study involves first-price auctions. In a first-price auction, poorly informed bidders do not benefit from the aforementioned informational linkage as in an English auction, explaining why [Carnehl and Weiergraeber \(2022\)](#)'s inefficiency is much stronger than ours.

<sup>45</sup>Column D corresponds actually to the probability of no-sale while a sale to a non-incumbent team would be efficient since the possibility of an efficient sale to the incumbent team is ruled out by our equilibrium analysis.

### 6.5.2 Reservation Wage Counterfactuals

Here we evaluate the effect of changing the reservation wage of a given player  $i$  on his expected revenue<sup>46</sup> and on his probability to be sold. For each  $i$ , we consider a baseline scenario and various counterfactual scenarios. In all scenarios, the variables  $(x_{i,f}, z_{i,f}, INC_{i,f})$  for all  $f$  take values as actually observed in the data. In the baseline scenario, the reservation wage also takes the observed value  $W_i^r$ , while in the counterfactual scenarios it either varies by -50 to +50 percent of the observed value or it is fixed to the sample minimum  $\min_{i=1,\dots,N} W_i^r$  (we have a total of 11 counterfactual scenarios). As above, 10,000 simulations are used to estimate the expected wages and sale probabilities under all scenarios.

Figure 1: Histogram of the Optimal Reserve



The figure shows the histograms of the optimal reserve (among the grid  $(5+k) \cdot 0.1 \cdot W_i^r$ ,  $k = 0, \dots, 10$ ) across the players  $i$ , by those who have an incumbent (*incumbent*) and those who do not (*no incumbent*). The optimal reserve corresponds to the reservation wage change which generates the maximum expected wage.

Before turning to the average impact (across players) of changing the reservation wage, we first compute for each player  $i$  the scenario that maximizes his expected wage. The reservation wage that maximizes expected wage is referred to as the optimal reserve. Figure 1 reports the distribution of the optimal reserve, first across the sub-sample of players with an incumbent and second across the sub-sample of players without an incumbent. We see that the lowest reservation wage (corresponding to the counterfactual scenario where the reserve is reduced by 50%) is optimal for about 40% of players in the sub-sample without an incumbent, and 30% in the sub-sample with an incumbent. The current reservation wage turns out to be optimal for 2 and 8% of players in the respective sub-samples. Note that a for a non-negligible fraction of players the optimal reserve is *above* the actually observed value.

Table 5 contains the average effects of modifying reservation wages. Column 1 reports, for each possible counterfactual scenario, the relative effect on the expected wage, and column 2 the effect on the probability of sale. More precisely, our wage effect is calculated by computing, for each  $i$ , the relative difference between the corresponding counterfactual expected wage and the expected wage in the baseline scenario, and by then taking the average over all players in

<sup>46</sup>Player  $i$ 's revenue is equal to his auction-determined wage  $w_i$  when he's sold, and zero otherwise.

the sample. Our sale probability effect is calculated similarly, i.e., we compute for each player the absolute difference in the estimated counterfactual and baseline probabilities to be sold, and then take the average over all players.<sup>47</sup> We emphasize that our counterfactual exercises do not consist in changing the reservation wages for *all players simultaneously* but rather in local policy experiments where we change the reservation wage of a given player  $i$  (while keeping it fixed for all others) and then calculate the resulting effect on the auction outcome for player  $i$  only (we do not analyze the effect this change may have on other players).<sup>48</sup> As Table 5 shows, changing the reservation wage generally has a substantial effect on both the expected wage and the probability to be sold. For instance, reducing the reserve by half is associated with a 13% increase in the expected wage and a 12-point percentage increase in the probability of being sold. Overall, we find that both the expected relative wage differential and the sale probability decrease monotonically with the reservation wage chosen in our counterfactual scenarios.

Table 5: Counterfactual results

| <i>Outcome of interest:</i>        | Final Wage Differential<br>(in relative terms) | Selling Probability Differential<br>(in absolute terms) |
|------------------------------------|--|---|
| <i>Counterfactual policy:</i>      |  |   |
| <b>Change in reservation wage:</b> |  |   |
| -50%                               | 13.22  | 0.12  |
| -40%                               | 7.27   | 0.09  |
| -30%                               | 5.83   | 0.06  |
| -20%                               | 4.32   | 0.04  |
| -10%                               | 0.93   | 0.02  |
| 10%                                | -3.94  | -0.02   |
| 20%                                | -2.02  | -0.03   |
| 30%                                | -1.28  | -0.04   |
| 40%                                | -4.33  | -0.06   |
| 50%                                | -4.15  | -0.07   |
| = sample minimum                   | 14.57  | 0.17  |

Note: Column 1 of the table reports the final wage differential (as percent different) between the counterfactual and baseline:  $\frac{1}{N} \sum_{i=1}^N \frac{\mathbb{E}(w_i \cdot \mathbf{1}\{i \text{ is sold}\} | \text{Counterfactual}) - \mathbb{E}(w_i \cdot \mathbf{1}\{i \text{ is sold}\} | \text{Baseline})}{\mathbb{E}(w_i \cdot \mathbf{1}\{i \text{ is sold}\} | \text{Baseline})}$ . Column 2 reports the difference in the probability of being sold between the counterfactual and baseline:  $\frac{1}{N} \sum_{i=1}^N [\mathbb{E}(\mathbf{1}\{i \text{ is sold}\} | \text{Counterfactual}) - \mathbb{E}(\mathbf{1}\{i \text{ is sold}\} | \text{Baseline})]$

The results of Table 5 together with those reported in Figure 1 suggest that, from the perspective of an individual player, lower reservation wages are beneficial. This contrasts with the empirical literature (starting with Paarsch (1997)) which typically finds that reserve prices actually observed in real-world auctions are substantially lower than the theoretically optimal ones. Individual deviations may of course be harmful for the group of players as a whole: e.g., if a wicket keeper  $i$  is getting hired by team  $f$  following a reduction in his reservation wage, then this lowers  $f$ 's valuations for the subsequent wicket keepers (through  $f$ 's auction shifters such as the backlog variable capturing the number of wicket-keepers previously bought and

<sup>47</sup>The exact formulas are given in the legend of Table 5.

<sup>48</sup>We prefer local counterfactual policies over global ones since the basic assumption required for correctly measuring a policy change –namely that all true parameter values are identical under both baseline and counterfactual configurations– is arguably less likely to hold in the latter case.

the remaining budget), which in turn should reduce the expected wage and sale probability of those wicket keepers that are sold after  $i$ . More generally, the fact that a reservation wage cut may be harmful to fellow players comes from the fact that the players being auctioned are partial substitutes. By setting reservation wages at a relative high level, the IPL avoids a prisoner's dilemma (whereby individual players would like to post a lower reservation wage), and appears to play a similar role as unions who negotiate for higher minimum wages in all sorts of branches (Neumark and Wascher (2007)).

### 6.5.3 Excluding the Incumbent

In our setup, excluding a non-incumbent reduces both the probability of sale and the worker's wage for any realization of the signals of all bidders. This reflects the fact that excluding a non-incumbent does not modify the bidding behavior of its opponents but reduces competition. On the contrary, excluding the incumbent has ambiguous effects because it modifies the bidding behavior of the non-incumbents. In particular, conditional on the incumbent's signal being such that its valuation is below the reservation wage, the cutoff wage of each non-incumbent is always lower when the incumbent is excluded than when the incumbent is allowed to participate. Some recent theoretical works have shown indeed that the exclusion of a bidder could enhance the seller's expected payoff (see, e.g., Bergemann et al. (2020) and Jehiel and Lamy (2020)).

We have therefore considered a counterfactual policy experiment where the incumbent is forbidden to participate in the auction. Expected wages and sale probabilities are estimated as above except that simulations under the counterfactual scenario are based on 7 instead of 8 firms (and calculations only concern the 75 auctions for which an incumbent without the RTM option is present). We find that excluding the incumbent reduces the expected wage by more than 60% while the sale probability is raised by about 10%. The drastic wage reduction occurs as in a private value model and is driven by the fact that the valuation of an incumbent is larger than the valuation of a non-incumbent *ceteris paribus* (since our estimate for  $\beta_{inc}^*$  is large). However, excluding a bidder would never lead to an increase of the sale probability in a private value environment, illustrating here the importance of the worker-specific signals: If the incumbent is allowed to participate but decides not to enter the auction, then this sends a bad signal to the non-incumbents who may be discouraged from entering the auction themselves.

## 7 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 interdependent value 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 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. We apply our methodology to a labor market application, analyzing how the auction-determined wages of sports players affect their performance, disentangling incentive and selection effects. As demonstrated by our application, our methodology also allows to develop rich counterfactual exercises.

While we develop our methodology using a labor market set-up and for a specific auction format – the English auction – our analysis can be easily adapted to other economic settings and other single-unit auction formats. The most promising alternative setting where our methodology can be applied is procurement auctions. Recent empirical works on procurement contracts awarded through competitive tendering have stressed the importance of both ex post moral hazard and adverse selection.<sup>49</sup> Our methodology can be applied more generally in settings wherein the remuneration rule that characterizes post-auction transfers depends on the auction winning price (i.e.,  $w_i$  in equation (1)) but also on verifiable variables that could be used as our post-auction outcome (i.e.,  $y_{i,f}^w$  in eq. (1)). If we take a quantity of outputs or inputs for the latter, then the auction shifters  $z_{i,f}$  could correspond to variables affecting the sunk costs of bidder  $f$  for the contract  $i$  while the characteristics  $x_{i,f}$  could capture variables affecting marginal costs (or benefits).

The extension of our approach to multi-unit or share auctions is also feasible but more challenging. For uniform and discriminatory auctions, we can apply results from econometrics of the multi-unit auctions (see [Hortaçsu \(2011\)](#) for a survey), in particular under pure private values. Multi-unit environments also open the door to novel post-auction specifications where not only the price but also the quantity purchased plays an incentive role and would allow to cover a broader class of applications, in particular involving the auctions organized by the Treasury or central bank. An active research question in macro-finance deals with the impact on macroeconomic outcomes of those auctions that are ubiquitous in the determination of the prices and quantities of key economic variables (for e.g., amount of debt, exchange rate, interest rates).<sup>50</sup> For example, [Joyce and Tong \(2012\)](#) study the impact of an increase in bond supply, through a quantitative easing program, on post-auction bond yields. Since both quantities and prices of bonds are determined by auctions, analyzing post-auction outcomes will be typically subject to both endogeneity and selection concerns. A range of applications in empirical corporate finance also examine post-acquisition outcomes of companies (e.g., firm sales) when they are acquired through an auction procedure (see [Eckbo \(2009\)](#) and [Dasgupta and Hansen \(2007\)](#) for reviews on auctions in corporate finance). For instance in bankruptcy

<sup>49</sup>E.g., [Lewis and Bajari \(2011\)](#) show that project-delivery time is driven by the contract characteristics both directly through incentives schemes, 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).

<sup>50</sup>Relatedly, [Cassola et al. \(2013\)](#) show that the regression of a bank's post-auction profitability on auction-determined 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). Note that [Cassola et al. \(2013\)](#) abstract from any incentive effect linking the auction and post-auction stages.

auctions, a question is whether recovery rates (the post-auction outcome  $y_{i,f_i^w}$  in eq. (1)) are higher when the previous owner of the bankrupt firm wins the auction (i.e., when  $f_i^w = f_i^{inc}$ ). The endogeneity and sample selection problems arise because previous owners tend to repurchase the firm when they share good news on the quality of the firm ( $s_i^{pws}$ ) and also because recovery rates are only observed for firms who have been successful in selling their assets.

**Acknowledgments:** We are grateful to the Editor (Elie Tamer) and three anonymous referees for many constructive comments and suggestions. We also 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, Umea University, the University of Franche-Comté, the University of Luxembourg, the University of Maryland, the University of Paris II, the 2nd ECODEC Conference, the 7th French Econometrics Conference, and the 46th EARIE conference, for helpful remarks, and to 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 prices on post-auction outcome.” Financial support from the LABEX ECODEC (ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047) and the ANR DACRERISK (Project-ANR-18-CE05-0004) is gratefully acknowledged.

## References

- ABITO, J. M. (2020): “Measuring the Welfare Gains from Optimal Incentive Regulation,” Review of Economic Studies, 87, 2019–2048.
- ABOWD, J. M., F. KRAMARZ, AND D. N. MARGOLIS (1999): “High Wage Workers and High Wage Firms,” Econometrica, 67, 251–333.
- ACEMOGLU, D. AND J.-S. PISCHKE (1998): “Why Do Firms Train? Theory and Evidence,” The Quarterly Journal of Economics, 113, 79–119.
- ACKERBERG, D. A., K. CAVES, AND G. FRAZER (2015): “Identification Properties of Recent Production Function Estimators,” Econometrica, 83, 2411–2451.
- AKERLOF, G. A. AND J. L. YELLEN (1990): “The Fair Wage-Effort Hypothesis and Unemployment,” Quarterly Journal of Economics, 105, 255–283.
- AKKUS, O., J. A. COOKSON, AND A. HORTAÇSU (2021): “Assortative Matching and Reputation in the Market for First Issues,” Management Science, 67, 2049–2074.
- 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.
- ASKER, J. AND E. CANTILLON (2008): “Properties of Scoring Auctions,” The RAND Journal of Economics, 39, 69–85.
- 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.
- ATHEY, S. AND J. LEVIN (2001): “Information and Competition in U.S. Forest Service Timber Auctions,” Journal of Political Economy, 109, 375–417.
- BERGEMANN, D., B. BROOKS, AND S. MORRIS (2020): “Countering the winner’s curse: Optimal auction design in a common value model,” Theoretical Economics, 15, 1399–1434.
- 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 (2022): “Bidding and Drilling Under Uncertainty: An Empirical Analysis of Contingent Payment Auctions,” Journal of Political Economy, 1319–1363.



- 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-College 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," *Quarterly Journal of Economics*, 134, 281–347.
- CARNEHL, C. AND S. WEIERGRAEBER (2022): "Bidder Asymmetries in Procurement Auctions: Efficiency vs. Information," *mimeo*.
- 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 Non-regular 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.
- DAS, M., W. NEWEY, AND F. VELLA (2003): "Nonparametric Estimation of Sample Selection Models," *Review of Economic Studies*, 70, 33–58.
- DASGUPTA, S. AND R. G. HANSEN (2007): "Auctions in corporate finance," in *Handbook of Empirical Corporate Finance*, Elsevier, 87–143.
- D'HAULTFOEUILLE, X. AND P. FÉVRIER (2020): "The Provision of Wage Incentives : A Structural Estimation Using Contracts Variation," *Quantitative Economics*, 11, 349–397.
- 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.
- ENGBRECHT-WIGGANS, R., P. R. MILGROM, AND R. J. WEBER (1983): "Competitive Bidding and Proprietary Information," *Journal of Mathematical Economics*, 11, 161–169.
- 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 (2000): "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68, 525–574.
- (2009): "Nonparametric Identification of Risk Aversion in First-Price Auctions Under Exclusion Restrictions," *Econometrica*, 77, 1193–1227.
- HAILE, P. AND E. TAMER (2003): "Inference with an Incomplete Model of English Auctions," *Journal of Political Economy*, 111, 1–51.
- 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., 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.
- 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 (2020): “On the Benefits of Set-Asides,” Journal of the European Economic Association, 18, 1655–1696.
- JOYCE, M. A. AND M. TONG (2012): “QE and the Gilt Market: a Disaggregated Analysis\*,” Economic Journal, 122, 348–384.
- KOMAROVA, T. (2017): “Extremum sieve estimation in k-out-of-n system.” Communications in Statistics - Theory and Method, 46, 4915–4931.
- KONG, Y. (2021): “Sequential Auctions with Synergy and Affiliation across Auctions,” Journal of Political Economy, 129, 148–181.
- KONG, Y., I. PERRIGNE, AND Q. VUONG (2022): “Multidimensional Auctions of Contracts: An Empirical Analysis,” American Economic Review, 112, 1703–36.
- 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.
- 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 (1986): Raids and Offer Matching, R. G. Ehrenberg, ed., Research in Labor Economics, vol. 8, 141–165.
- (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,” 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.
- 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.
- NEUMARK, D. AND W. WASCHER (2007): “Minimum Wages and Employment,” Foundations and Trends(R) in Microeconomics, 3, 1–182.
- OLLEY, G. S. AND A. PAKES (1996): “The Dynamics of Productivity in the Telecommunications Equipment Industry,” Econometrica, 64, 1263–1297.

- PAARSCH, H. J. (1997): "Deriving an Estimate of the Optimal Reserve Price: An Application to British Columbian Timber Sales," Journal of Econometrics, 78, 333–357.
- PAARSCH, H. J. AND H. HONG (2006): An Introduction to the Structural Econometrics of Auction Data, MIT Press.
- PERRIGNE, I. AND Q. VUONG (2011): "Nonparametric Identification of a Contract Model With Adverse Selection and Moral Hazard," Econometrica, 79, 1499–1539.
- (2021): "Econometric for Auctions," in preprint for the Handbook of Econometrics, Vol. 7B, Elsevier, London.
- PINKSTON, J. (2009): "A Model of Asymmetric Employer Learning with Testable Implications," Review of Economic Studies, 76, 367–394.
- POWELL, D. AND D. GOLDMAN (2021): "Disentangling Moral Hazard and Adverse Selection in Private Health Insurance," Journal of Econometrics, 222, 141–160.
- ROBERTS, J. W. (2013): "Unobserved heterogeneity and reserve prices in auctions," RAND Journal of Economics, 44, 712–732.
- RYAN, N. (2020): "Contract Enforcement and Productive Efficiency: Evidence from the Bidding and Renegotiation of Power Contracts in India," Econometrica, 88, 383–424.
- SHAPIRO, C. AND J. E. STIGLITZ (1984): "Equilibrium unemployment as a worker discipline device," American Economic Review, 74, 433–444.
- SHEARER, B. (2004): "Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment," Review of Economic Studies, 71, 513–534.
- VELLA, F. (1998): "Estimating Models with Sample Selection Bias: A Survey," Journal of Human Resources, 33, 127–169.
- 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.

# Appendices A-C

## A Equilibrium analysis

In this appendix we formally derive the equilibrium strategy associated with the auction model of Section 3. As a preliminary, let us present and discuss three additional technical assumptions that are needed for the analysis.

**TA1:** For each auction  $i$  with an incumbent, the function  $\kappa_i : \mathbb{R}_+ \rightarrow \mathbb{R}$ , defined by  $\kappa_i(u) := \log(\psi(u)) - \log(\mathbb{E}[e^{\lambda \cdot [s_i^{I,ws} + s_i^{P,ws}]} | w_{i,f_i}^* = u])$ , is an increasing bijection.

**TA2:** For each  $i$  and  $f = 1, \dots, F$ , the commonly known CDF of  $f$ 's aggregate signal  $\bar{s}_{i,f}$  conditional on  $(x_{i,f}, z_{i,f})_{f=1, \dots, F}$  and  $f_i^{inc}$ , denoted  $G_{i,f}$ , is atomless and has full support on  $\mathbb{R}$ .

**TA3:** In addition to the publicly observable covariates  $(x_{i,f}, z_{i,f}, INC_{i,f})_{f=1, \dots, F}$ , firms' payoff functions and the joint distribution of the private signals are commonly known by all firms.

Given A1 ii), TA1 is equivalent to stating that the difference  $\log(\psi(u)) - \log(\mathbb{E}[\psi(w_{i,f}^*) | w_{i,f_i}^* = u])$  is increasing in  $u$ .<sup>51</sup> Assumption TA1 imposes thus a restriction on the relationship between the incumbent's valuation  $w_{i,f_i}^*$  and the valuation of any given non-incumbent  $f$ : it says that the percentage increase of  $\psi(w_{i,f_i}^*)$  is larger than the percentage increase of  $\mathbb{E}[\psi(w_{i,f}^*) | w_{i,f_i}^* = u]$  when the incumbent's valuation increases. TA1 is analogous to the kind of assumptions typically made in the literature on auctions with interdependent values (see, e.g., the "average crossing property" in Krishna (2002)). As shown in the SA, TA1 holds when aggregate signals follow Gaussian distributions. Assumption TA2 guarantees that the incumbent's distribution has full support which ensures equilibrium uniqueness. Regarding equilibrium uniqueness, note that Bikhchandani et al. (2002) has shown that the symmetric model of Milgrom and Weber (1982) involves a continuum of perfect Bayesian equilibria in weakly undominated strategies, except in the pure private value case where such an equilibrium is unique. Here we establish uniqueness in an asymmetric model. TA3 is a type of assumption usually made in the literature (explicitly or often implicitly). It implies in particular that the scalars  $\bar{V}_{i,f}$ ,  $\lambda$  and the function  $\tau(\cdot)$  are common knowledge.

We now turn to the equilibrium characterization. Given A1 i), the auctions for the various workers  $i = 1, \dots, N$  can be analyzed separately from each other. We can thus focus the analysis on one arbitrary auction, and for notational simplicity the index  $i$  is suppressed in the remainder of this appendix.

Let us define the function  $w_{f^{inc}} : \mathbb{R} \rightarrow \mathbb{R}_+$  by  $w_{f^{inc}}(x) := \psi^{-1}(e^{\lambda \cdot [\bar{V}_{f^{inc}} + x]})$ . Under A2 and since  $\lambda > 0$ , this function is an increasing bijection. We then use the notation  $w_{f^{inc}}^{-1}$  to denote its inverse, so that  $w_{f^{inc}}^{-1}(u) = (\log(\psi(u))/\lambda) - \bar{V}_{f^{inc}}$ . The incumbent's payoff function when winning at the wage  $w$  takes the form  $e^{\lambda \cdot [\bar{V}_{f^{inc}} + \bar{s}_{f^{inc}} + \tau(w)]} - w$ . Given A2, this payoff multiplied by the strictly positive term  $e^{-\lambda \cdot \tau(w)}$  is a decreasing function. As a consequence, the payoff of the incumbent crosses zero only at the cutoff wage  $w_{f^{inc}}(\bar{s}_{f^{inc}}) = w_{f^{inc}}^*$ , and is

<sup>51</sup>Formally, we use that  $\mathbb{E}[\psi(w_{i,f}^*) | w_{i,f_i}^* = u] = \mathbb{E}[e^{\lambda \cdot [\bar{V}_{i,f} + \bar{s}_{i,f} + s_i^{I,ws} + s_i^{P,ws}]} | w_{i,f_i}^* = u] = \mathbb{E}[e^{\lambda \cdot [\bar{V}_{i,f} + \bar{s}_{i,f}]}] \cdot \mathbb{E}[e^{\lambda \cdot [s_i^{I,ws} + s_i^{P,ws}]} | w_{i,f_i}^* = u]$ , where the last equality comes the fact that, under A1 ii), the signals  $\bar{s}_{i,f}$  and  $\bar{s}_{i,f_i}^{inc}$  are independent. We obtain then that the difference between  $\log(\psi(u)) - \log(\mathbb{E}[\psi(w_{i,f}^*) | w_{i,f_i}^* = u])$  and  $\kappa_i(u)$  is equal to  $\log(\mathbb{E}[e^{\lambda \cdot [\bar{V}_{i,f} + \bar{s}_{i,f}]}])$ , which does not depend on  $u$ .

positive (resp. negative) when the wage is below (resp. above) this cutoff wage. Therefore, a weakly dominant strategy for the incumbent consists in entering the auction only if  $w_{finc}^* \geq W^r$ , and, when this is so, in remaining active until the clock reaches the valuation  $w_{finc}^*$ . Given our full support TA2, we can show that this strategy is actually the unique best-response for the incumbent: if this firm adopts another strategy, it would raise a strictly lower expected payoff. For any alternative strategy for the incumbent, we can develop exactly the same analysis as below which establishes in particular that there is a positive probability to win the auction if the incumbent bids slightly above its valuation.

To determine the best responses of non-incumbents, we need to distinguish four cases. These four cases differ in the incumbent's bidding history. Case A (resp. B) corresponds to a bidding history wherein the incumbent is not present (resp. is present but chooses not to enter the auction). Cases C and D consider bidding histories wherein the incumbent did participate. C covers the case where the incumbent has decided to drop out prior to the current value of the auction clock, while D the case where the incumbent is still active at the clock's current value.

In cases A, B, and C, the beliefs of non-incumbents regarding the sum of worker-specific signals  $s_i^{I,ws} + s_i^{P,ws}$  do not evolve but stay fixed during (the remainder of) the auction: This property results from A1 ii) which guarantees that the non-incumbents share the same common beliefs which are determined by Bayesian updating given that the incumbent follows its weakly undominated strategy. The payoff function of  $f$  then takes the form

$$e^{\lambda \cdot [\bar{V}_f + \bar{s}_f + \tau(w)]} \cdot H - w \quad (\text{A.1})$$

where  $H$  represents the expectation of  $e^{\lambda \cdot [s_i^{I,ws} + s_i^{P,ws}]}$  given the incumbent's bidding history. This expectation thus differs across the cases A, B and C, but under A1 ii) it does not depend on  $s_{i,f}$ . Given A2, for the same reason as above, the expected payoff of the non-incumbent firm  $f$  (its beliefs being captured by the constant  $H$ ) crosses zero only at the cutoff wage  $\psi^{-1}(e^{\lambda \cdot [\bar{V}_{i,f} + \bar{s}_{i,f}]} \cdot H)$ . Below we detail the precise form of  $H$  for each of the three cases together with the optimal strategies. In case D, the incumbent is still actively bidding and the equilibrium strategy of non-incumbents is obtained in a similar manner as in [Milgrom and Weber \(1982\)](#).

**Case A: there is no incumbent.** The constant  $H$  in (A.1) is equal to the unconditional expectation  $\mathbb{E}[e^{\lambda \cdot [s_i^{I,ws} + s_i^{P,ws}]}]$ . Let  $w_f[NI](\bar{s}_f) := \psi^{-1}(e^{\lambda \cdot [\bar{V}_f + \bar{s}_f]} \cdot \mathbb{E}[e^{\lambda \cdot [s_i^{I,ws} + s_i^{P,ws}]}])$  denote the corresponding cutoff wage, where the acronym *NI* stands for No Incumbent. It is then a best-response for  $f$  to enter the auction only if  $w_f[NI](\bar{s}_f) \geq W^r$ , and then to remain active until the clock has reached this cutoff wage otherwise. Under A2 and since  $\lambda > 0$ , the function  $w_f[NI] : \mathbb{R} \rightarrow \mathbb{R}_+$  is an increasing bijection.

**Case B: the incumbent has decided not to enter the auction.** From Bayesian updating given the incumbent's equilibrium strategy, non-incumbents can then infer that  $w_{finc}^* < W^r$  such that the constant  $H$  in (A.1) is equal to  $\mathbb{E}[e^{\lambda \cdot [s_i^{I,ws} + s_i^{P,ws}]} | \bar{s}_{finc} < w_{finc}^{-1}(W^r)]$ . Let  $w_f[NE](\bar{s}_f) = \psi^{-1}(e^{\lambda \cdot [\bar{V}_f + \bar{s}_f]} \cdot H)$  denote the corresponding cutoff wage, where the acronym *NE* stands for No Entry. It is then a best-response for  $f$  to enter the auction only if  $w_f[NE](\bar{s}_f) \geq W^r$ , and then to remain active until the clock has reached this cutoff wage otherwise. Under A2 and since  $\lambda > 0$ , the function  $w_f[NE] : \mathbb{R} \rightarrow \mathbb{R}_+$  is an increasing bijection.

**Case C: the incumbent has entered the auction and dropped out.** Suppose that the current value of the auction clock is  $\tilde{w}$ . In this case, the incumbent has necessarily dropped out from the auction at  $w' \in [W^r, \tilde{w})$ . From Bayesian updating and given the incumbent's equilibrium strategy, non-incumbents can then infer that  $\bar{s}_{finc} = w_{finc}^{-1}(w')$  and the constant  $H$  in (A.1) is

thus equal to  $\mathbb{E}[e^{\lambda \cdot [s^{I,ws} + s^{P,ws}]} | \bar{s}_{f^{inc}} = w_{f^{inc}}^{-1}(w')]$ . Let  $w_f[DROP](\bar{s}_f; w') = \psi^{-1}(e^{\lambda \cdot [\bar{V}_f + \bar{s}_f]} \cdot H)$  denote the corresponding cutoff wage. It is then a best-response for the non-incumbent  $f$  to drop out immediately if this cutoff wage stands below the current wage  $\tilde{w}$ , and to remain active until the clock has reached the cutoff wage otherwise. Under A2 and since  $\lambda > 0$ , the function  $w_f[DROP](\cdot; w') : \mathbb{R} \rightarrow \mathbb{R}_+$  is an increasing bijection.

So far we have studied situations where the incumbent is either absent or left the auction at some point, implying that the incumbent's bidding history is fixed, which in turn means that beliefs of non-incumbents regarding the sum of worker-specific signals  $s^{I,ws} + s^{P,ws}$  do not evolve during the remainder of the auction. In those cases, the bidding incentives are the same as to the ones under a pure private value environment and the equilibrium best-responses are unique given the full support assumption TA2. Now we consider the remaining case where the incumbent is still active in the auction and we establish that non-incumbents have a unique best-response.

**Case D: the incumbent has entered the auction and is still active.** Let  $w_f[ACT](\bar{s}_f)$  be the wage such that  $f$ 's expected payoff is null conditional on winning the auction at  $w_f[ACT](\bar{s}_f)$ , and assuming that  $f^{inc}$  drops out exactly at this value, from which we can infer that the incumbent's aggregate signal is equal to  $w_{f^{inc}}^{-1}(w_f[ACT](\bar{s}_f))$ . The wage  $w_f[ACT](\bar{s}_f)$  depends on  $f$ 's vector of signals only through its aggregate signal and is characterized as the solution, in  $u$ , of the equation  $e^{\lambda \cdot [\bar{V}_f + \bar{s}_f]} \cdot \mathbb{E}[e^{\lambda \cdot [s^{I,ws} + s^{P,ws}]} | w_{f^{inc}}^* = u] = \psi(u)$ . Using the definition of  $\kappa(\cdot)$ , this equation is equivalent to  $\kappa(u) = \lambda \cdot [\bar{V}_f + \bar{s}_f]$  and has a solution in  $u$  which is unique given that  $\kappa$  is assumed to be a bijection (TA1). Furthermore, the function  $w_f[ACT] : \mathbb{R} \rightarrow \mathbb{R}_+$  is an increasing bijection since we have assumed that  $\kappa$  is also an increasing bijection and  $\lambda > 0$ . We show below that  $f$ 's unique best response is as follows: as long as  $f^{inc}$  is active then  $f$  should also remain active at the clock value  $\tilde{w}$  if and only if  $w_f[ACT](\bar{s}_f)$  exceeds  $\tilde{w}$ ; once  $f^{inc}$  has dropped out from the auction, the non-incumbent  $f$  should from that moment onward adopt the strategy associated with case C.

For any  $u \geq 0$ , let  $\Pi_f[f^{inc}](\bar{s}_f, u)$  denote the non-incumbent firm  $f$ 's expected payoff conditional on the incumbent  $f^{inc}$  having entered the auction, and when it uses the following strategy: remain active as long as the incumbent is active and provided the clock has not reached the value  $u$ ; drop out at  $u$  if  $f^{inc}$  is still active at this value; if  $f^{inc}$  drops out before  $u$ , then immediately switch to the strategy of case C from that drop-out value onward. The expectation, calculated at the start of the auction, is with respect to the drop-out value of  $f^{inc}$ . To detail the expression of  $\Pi_f[f^{inc}](\bar{s}_f, u)$ , let  $p_f(x)$  denote the belief of firm  $f$  about the probability that the auction stops at  $x$  if the incumbent  $f^{inc}$  drops out at  $x \geq W^r$  (i.e., the probability that all possible other competitors of  $f$  have dropped out before or exactly at  $x$ ). From case C we know that if the incumbent dropped out at  $x$ , then any non-incumbent  $f'$  with an aggregate signal below  $[w_f[DROP]]^{-1}(x; x)$  should have dropped out before  $x$  or immediately at this value. We thus have  $p_f(x) \geq \prod_{f' \neq f, f^{inc}} G_{f'}([w_f[DROP]]^{-1}(x; x))$  where the function  $[w_f[DROP]]^{-1}(\cdot; x)$  denotes the inverse of the function  $y \rightarrow w_f[DROP](y; x)$ . Thanks to our full support assumption on the aggregate signals (TA2), the above product is strictly positive and we have thus  $p_f(x) > 0$  for any  $x \geq W^r$ . As we will see, the explicit form of  $p_f(x)$  (which depends on the number of other non-incumbents and their exact strategies) does not play any role in characterizing  $f$ 's optimal bidding strategy. However, the fact that this probability is strictly positive guarantees that the equilibrium best-response is unique as shown below. We also let  $\Pi_f^{cont}(\bar{s}_f, x)$  denote the continuation payoff of  $f$  if the incumbent has dropped out at  $x \geq W^r$  and if there is at least one other non-incumbent active at this value. Finally, for any

$x \geq W^r$ , we also let

$$\Pi_f(\bar{s}_f, x) = \left[ \mathbb{E}[e^{\lambda \cdot [s^{L,ws} + s^{P,ws}]} | w_{f^{inc}}^* = x] \cdot e^{\lambda \cdot [\bar{V}_f + \bar{s}_f + \tau(x)]} - x \right] \cdot p_f^{f^{inc}}(x) + \Pi_f^{cont}(\bar{s}_f, x) \cdot (1 - p_f^{f^{inc}}(x)). \quad (\text{A.2})$$

The term  $\Pi_f(\bar{s}_f, x)$  corresponds to the expected payoff of  $f$ , calculated when the incumbent quit the auction at  $x$ , and if  $f$  is still active at  $x$  and follows thereafter the (optimal) strategy associated to case C.

We have then

$$\Pi_f[f^{inc}](\bar{s}_f, u) = \int_{w_{f^{inc}}^{-1}(W^r)}^{w_{f^{inc}}^{-1}(u)} \Pi_f(\bar{s}_f, w_{f^{inc}}^*(\bar{s}_{f^{inc}})) \cdot \frac{dG_{f^{inc}}(\bar{s}_{f^{inc}})}{1 - G_{f^{inc}}(w_{f^{inc}}^{-1}(W^r))}. \quad (\text{A.3})$$

The lower limit of the integral in (A.3) follows from the fact that the incumbent is assumed to have entered the auction, and the upper limit from the fact that  $f$ 's payoff equals zero if the incumbent quits the auction after this firm, that is to say drops out after  $u$ . In order to show now that  $\text{Arg max}_{u \geq W^r} \Pi_f[f^{inc}](\bar{s}_f, u) = w_f[ACT](\bar{s}_f)$ , we establish below that  $\Pi_f(\bar{s}_f, w_f[ACT](\bar{s}_f)) = 0$ ,  $\Pi_f(\bar{s}_f, u) > 0$  for  $u < w_f[ACT](\bar{s}_f)$ , and  $\Pi_f(\bar{s}_f, u) < 0$  for  $u > w_f[ACT](\bar{s}_f)$ .

By definition of the function  $w_f[ACT](\cdot)$ , the term within brackets in (A.2) is null if  $x = w_f[ACT](\bar{s}_f)$ . Furthermore, if the incumbent drops out exactly at  $w_f[ACT](\bar{s}_f)$ , then firm  $f$ 's best continuation strategy is then to drop out immediately which yields a null payoff (the payoff is null either if  $f$  loses the auction or if  $f$  wins at  $w_f[ACT](\bar{s}_f)$ ). Thus  $\Pi_f^{cont}(\bar{s}_f, w_f[ACT](\bar{s}_f)) = 0$ , and on the whole we have  $\Pi_f(\bar{s}_f, w_f[ACT](\bar{s}_f)) = 0$ . Next we check that  $\Pi_f(\bar{s}_f, u')$  is larger than zero for all  $u' < w_f[ACT](\bar{s}_f)$ . Since  $w_f[ACT]$  is an increasing bijection, there exists an aggregate signal  $\bar{s}'_f < \bar{s}_f$  such that  $u' = w_f[ACT](\bar{s}'_f)$ . From the definition of  $w_f[ACT]$ , it follows that  $\mathbb{E}[e^{\lambda \cdot [s^{L,ws} + s^{P,ws}]} | w_{f^{inc}}^* = u'] \cdot e^{\lambda \cdot [\bar{V}_f + \bar{s}'_f + \tau(u')]} = u'$ . Given that  $\lambda > 0$  and  $\bar{s}'_f < \bar{s}_f$ , we have thus  $\mathbb{E}[e^{\lambda \cdot [s^{L,ws} + s^{P,ws}]} | w_{f^{inc}}^* = u'] \cdot e^{\lambda \cdot [\bar{V}_f + \bar{s}_f + \tau(u')]} - u' > 0$ . We have also  $\Pi_f^{cont}(\bar{s}_f, u') \geq 0$  since firm  $f$  would guarantee itself a positive payoff by dropping out immediately at the wage  $u'$ . Since  $p_f(u') > 0$ , we obtain finally that  $\Pi_f(\bar{s}_f, u') > 0$  for any  $u' < w_f[ACT](\bar{s}_f)$ . Similarly, we obtain that  $\Pi_f(\bar{s}_f, u') < 0$  for any  $u' > w_f[ACT](\bar{s}_f)$ .

Hence we have established that  $\text{Arg max}_{u \geq W^r} \Pi_f[f^{inc}](\bar{s}_f, u) = \{w_f[ACT](\bar{s}_f)\}$ . Therefore, as long as the incumbent is active in the auction, each non-incumbent  $f$  has a unique optimal strategy (independently of the way the other non-incumbents are bidding) which consists in staying in the auction until the cutoff wage  $w_f[ACT](\bar{s}_f)$ .

The following proposition summarizes our equilibrium characterization and the key properties we use for identification.

**Proposition A.1.** *Under A1, A2 and TA1-TA3, there is a unique perfect Bayesian equilibrium. The strategy of the incumbent consists in not entering the auction if  $w_{f^{inc}}(\bar{s}_{f^{inc}})$  is below  $W^r$ , and in remaining active until  $w_{f^{inc}}(\bar{s}_{f^{inc}})$  otherwise. The strategy of a non-incumbent firm  $f$  depends on its aggregate signal  $\bar{s}_f$  and on the incumbent's bidding history as follows:*

- *Participation decision of non-incumbents: If there is no incumbent then firm  $f \neq f^{inc}$  should enter only if  $w_f[NI](\bar{s}_f) \geq W^r$ . If instead an incumbent is present and this firm entered (resp. did not enter) the auction, then  $f$  should enter only if  $w_f[DROP](\bar{s}_f; W^r) \geq W^r$  (resp.  $w_f[NE](\bar{s}_f) \geq W^r$ ).*
- *Drop-out decision of non-incumbents: Suppose firm  $f$  entered the auction and the auction clock has reached  $\tilde{w}$ . If there is no incumbent (resp. the incumbent has not entered the*

auction), then  $f$  should exit instantly if  $w_f[NI](\bar{s}_f) < \tilde{w}$  (resp.  $w_f[NE](\bar{s}_f) < \tilde{w}$ ), and should remain active otherwise; if the incumbent has dropped out at  $w_{f^{inc}}^* < \tilde{w}$  then  $f$  should exit instantly if  $w_f[DROP](\bar{s}_f; w_{f^{inc}}^*) < \tilde{w}$ , and remain active otherwise; finally, if the incumbent is still active at  $\tilde{w}$ , then  $f$  should exit instantly if  $w_f[ACT](\bar{s}_f) < \tilde{w}$ , and remain active otherwise.

The cutoff wage functions  $w_{f^{inc}}(\cdot)$ ,  $w_f[NI](\cdot)$ ,  $w_f[NE](\cdot)$ ,  $w_f[DROP](\cdot; x)$  (for any  $x \geq W^r$ ) and  $w_f[ACT](\cdot)$  are increasing bijections from  $\mathbb{R}$  to  $\mathbb{R}_+$ .

The equilibrium strategies are such that, for any realization of bidders' private signals, we can recover in a deterministic way, the equilibrium outcome, i.e., whether the worker is sold and if so, the identity of the winning firm and the wage paid to this worker. Note that our model is thus a complete model in contrast to Haile and Tamer (2003)'s incomplete model.

A final result (used to establish the identification of our bidding model) concerns the way in which firms' equilibrium strategies depend on their private signals and the model primitives. The equilibrium cutoff wage  $w_{f^{inc}}(\bar{s}_{f^{inc}})$  for an incumbent firm is equal to  $\psi^{-1}(\lambda \cdot (\bar{V}_{f^{inc}} + \bar{s}_{f^{inc}}))$ . This expression depends solely on  $\lambda \cdot \bar{s}_{f^{inc}}$  and two model primitives:  $\lambda \cdot \bar{V}_{f^{inc}}$  and the function  $\psi$ . The equilibrium strategy of a non-incumbent firm  $f$  depends on whether there is an incumbent and if so on what has been observed so far from the incumbent's bidding history. In case A, the expression of the cutoff wage  $w_f[NI](\bar{s}_f)$  for the non-incumbent firm  $f$  depends solely on  $\lambda \cdot \bar{s}_f$  and three model primitives:  $\lambda \cdot \bar{V}_f$ , the function  $\psi$  and the distribution of  $\lambda \cdot (s^{I,ws} + s^{P,ws})$ . In cases B, C and D, the expression of the corresponding cutoff wages of the non-incumbent firm  $f$  depends solely on  $\lambda \cdot \bar{s}_f$ ,  $W^r$ ,<sup>52</sup>  $\lambda \cdot \bar{V}_f$ , the function  $\psi$  and the distribution of  $\lambda \cdot (s^{I,ws} + s^{P,ws})$  conditional on the incumbent's bidding history (this conditional distribution allows us to compute the term  $H$  in the expressions above). Given that the incumbent uses the cutoff wage  $w_{f^{inc}}(\bar{s}_{f^{inc}})$ , and thanks to the independence assumption A1 ii), the distribution of  $\lambda \cdot (s^{I,ws} + s^{P,ws})$  conditional on the incumbent's bidding history can be expressed solely in terms of  $\lambda \cdot \bar{V}_{f^{inc}}$ ,  $\psi$ , and the (unconditional) distributions of  $\lambda \cdot (s^{I,ws} + s^{P,ws})$  and  $\lambda \cdot (s_{f^{inc}}^{I,ms} + s_{f^{inc}}^{P,ms})$ .

The following corollary summarizes the discussion above: it states which primitives of our model need to be known in order to identify the equilibrium strategy of each firm, and hence the distribution of the bidding history.

**Corollary 1.** *The knowledge of the model primitives  $\{\lambda \cdot \bar{V}_f\}_{f=1,\dots,F}$ , the function  $\psi$ , and the distributions of both  $\lambda \cdot (s_f^{I,ms} + s_f^{P,ms})$  (for each  $f$ ) and  $\lambda \cdot (s^{I,ws} + s^{P,ws})$ , is sufficient to identify each firm's equilibrium strategy as a function of  $\lambda \cdot \bar{s}_f$  for any reservation wage  $W^r$ , and hence the distribution of the bidding history.*

An implication of this corollary is that we do not need to identify the parameter  $\lambda$  in order to identify the distribution of the bidding history and then the control terms that are included (in our second step) into the performance equation. This is crucial for our two-stage identification strategy since  $\lambda$  is not identified from the bidding data only.

## B Proof of Proposition 4.1

Let us first present formally two additional technical assumptions that complete the assumptions needed for Proposition 4.1.

**TA4:** i) For any given vector of covariates  $COV \in \mathcal{S}_{COV}$  and any given firm  $f = 1, \dots, F$ , then conditional on  $COV_i = COV$ , there exists a set of potential participants  $S$  with  $f \in S$  and

<sup>52</sup>The dependence on  $W^r$  arises only in case A.



$|S| \geq 2$  such that we have with positive probability  $S_i = S$ ,  $f \in S_i$  and  $\sum_{f' \in S_i} INC_{i,f'} = 0$ ; ii) There exists a pair of firms  $(\tilde{f}, \tilde{f}')$  such that for any vector of covariates  $COV \in \mathcal{S}_{COV}$ , then conditional on  $COV_i = COV$ , there is a positive probability that  $f_i^{inc} = \tilde{f}$  and  $S_i = \{\tilde{f}, \tilde{f}'\}$ .

TA4 means that conditional on a given vector of covariates, we observe both auctions with and without an incumbent. Furthermore, TA4 i) means that in the sub-sample of auctions without an incumbent, each firm  $f$  participates with positive probability. TA4 ii) means there is a positive probability that the auction involves only two firms one of which being an incumbent. This assumption allows to identify  $H_{\tilde{f}'}^{ACT}(\cdot | COV, \tilde{f})$  and  $H_{\tilde{f}}^{inc}(\cdot | COV, \tilde{f})$  for any  $COV \in \mathcal{S}_{COV}$  for a given pair  $(\tilde{f}, \tilde{f}')$ .

In assumption A6, we consider a function that depends on two firms  $\tilde{f}, \tilde{f}' \in \{1, \dots, F\}$ , and for any such pair we assume that the given function is injective. Indeed, we need this function to be injective only for one pair: a pair  $(\tilde{f}, \tilde{f}')$  that guarantees that TA4 ii) hold.

**TA5:** For any firm  $f \in \{1, \dots, F\}$  and any set of covariates  $COV = (\mathbf{x}_{f'}, \mathbf{z}_{f'})_{f'=1, \dots, F} \in \mathcal{S}_{COV}$ , then there exists  $COV^* = (\mathbf{x}_{f'}^*, \mathbf{z}_{f'}^*)_{f'=1, \dots, F} \in \mathcal{S}_{COV}$  such that  $(\mathbf{x}_{\tilde{f}}^*, \mathbf{z}_{\tilde{f}}^*) = (\mathbf{x}_f, \mathbf{z}_f)$  where the firm  $\tilde{f}$  has been defined in TA4 ii).

In words, we need that the set of covariates that the special firm  $\tilde{f}$  can have is larger than the set of covariate than the other firms can have. E.g. TA5 hold if the set  $\mathcal{S}_{COV}$  is symmetric to permutation, in which cases those sets are equal.

Let us first identify the distributions  $H_{\tilde{f}}^{inc}(\cdot | COV, \tilde{f})$ ,  $\{H_f^{NI}(\cdot | COV)\}_{f=1, \dots, F}$  and  $H_{\tilde{f}'}^{ACT}(\cdot | COV, \tilde{f})$  for any  $COV \in \mathcal{S}_{COV}$ . Take a given  $COV \in \mathcal{S}_{COV}$  and a given firm  $f^* \in \{1, \dots, F\}$ . TA4 i) guarantees that conditional on  $COV \in \mathcal{S}_{COV}$ , there exists a set of potential participants  $S_{f^*}$  that contains at least two bidders including  $f^*$  and such that  $S_i = S_{f^*}$  and  $f_i^{inc} \notin S_{f^*}$  with a strictly positive probability. Consider first the sub-sample of workers  $i$  for which the set of potential participants is equal to the set  $S_{f^*}$ . Conditional on  $COV$  and when there is no incumbent, our bidding model is observationally equivalent to an independent pure private value model: thanks to A1 ii), the cutoff wages  $w_{i,f}[NI](\bar{s}_{i,f}) = \psi^{-1}\left(e^{\lambda \cdot [\bar{V}_{i,f} + \bar{s}_{i,f}]} \cdot \mathbb{E}[e^{\lambda \cdot (s_i^{J,ws} + s_i^{B,ws})} | COV_i = COV]\right)$  are drawn independently across the bidders  $f \in S_i$ . We know from [Athey and Haile \(2002\)](#) that the independent asymmetric private values model is non-parametrically identified under the English button auction from the joint observation of the identity of the winner and the winning price if there are at least two bidders, no reserve price and if the supports of the (atomless) distributions of bidders' private values are the same for all bidders.<sup>53</sup> In our setup, TA2 guarantees that the distribution of the cutoff wages  $w_{i,f}[NI](\bar{s}_{i,f})$  (for  $f \in S_{f^*}$ ) are atomless and their supports is  $\mathbb{R}_+$  for all firms. Transposing this result to our setting (and the sub-sample where the set of participants is  $S_{f^*}$ ), we identify the distribution  $H_{f^*}^{NI}(\cdot | COV)$ . The argument works for any  $f^* \in \{1, \dots, F\}$  and hence we identify  $\{H_f^{NI}(\cdot | COV)\}_{f=1, \dots, F}$ .

Second, we consider the sub-sample of workers  $i$  for which the set of potential participants  $S_i$  corresponds to  $(\tilde{f}, \tilde{f}')$  and where  $\tilde{f}$  is the incumbent (i.e.  $f_i^{inc} = \tilde{f}$ ). From TA4 ii), this arises with positive probability conditional on any  $COV \in \mathcal{S}_{COV}$ . In this sub-sample, the bidding model is also observationally equivalent to an independent pure private value model: it is as if the incumbent  $\tilde{f}$  and the non-incumbent  $\tilde{f}'$  have the private values  $w_{i,\tilde{f}}^*$  and  $w_{\tilde{f}'}[ACT](\bar{s}_{i,\tilde{f}'} | COV_i, \tilde{f})$ , respectively, where for each firm  $f$ ,  $w_f[ACT](\bar{s}_{i,f} | COV_i, f_i^{inc})$  denotes

<sup>53</sup>Formally, identification hinges on the following mathematical 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. Note that the atomless assumption allows us to abstract from the way ties are solved.

the cutoff wage which is the solution of (5). Applying again Athey and Haile (2002) allows to identify the distributions  $H_{\tilde{f}}^{inc}(\cdot|COV, \tilde{f})$  and  $H_{\tilde{f}}^{ACT}(\cdot|COV, \tilde{f})$ .

**Remark:** In auctions with an incumbent and strictly more than two bidders, the bidding data is no longer observationally equivalent to a pure private value model, i.e. to a model where the strategy of each bidder reduces to a single cutoff wage up to which it remains active which prevents to apply Athey and Haile (2002) if only the final wage and the identity of the winner of observed. However, if the bidding behavior of the incumbent can be also observed by the econometrician, then we do not longer need TA4 ii): it is sufficient to rely on the final step of the auctions when there are only two active bidders one being the incumbent to identify the bidding model. Instead of A5 and TA4 we could simply assume that the set of participants is fixed, contains at least two firms and that both auctions with and without an incumbent occur with positive probability.

Then A6 allows us to identify the function  $\psi$ . At the end of this Appendix we will show that the condition A6 is automatically satisfied if we parameterize the function  $\tau$  such that  $\tau(w) = \tau \cdot \log(w)$ .

Let us define the function  $\bar{\psi} : \mathbb{R}_+ \rightarrow \mathbb{R}$  by  $\bar{\psi}(x) := \log(\psi(x))$ . From A2, the function  $\bar{\psi}$  is an increasing bijection and let  $\bar{\psi}^{-1}$  denote its inverse function. Naturally  $\bar{\psi}^{-1}$  is identified since  $\psi$  is identified. This allows us to identify the distribution of  $\lambda \cdot (\bar{V}_{i,f} + s_{i,f}^{I,ms} + s_f^{P,ms} + s_i^{I,ws} + s_i^{P,ws}) = \bar{\psi}(w_{i,f}^*)$  for  $f = \tilde{f}$ : this distribution is equal to  $H_{\tilde{f}}^{inc}(\bar{\psi}^{-1}(\cdot)|COV, \tilde{f})$ . From the normalization A1 iii), we thus identify the scalar  $\lambda \cdot \bar{V}_{i,\tilde{f}}$  (for the workers  $i$  such that  $COV_i = COV$  and  $f_i^{inc} = \tilde{f}$ )<sup>54</sup> and the distribution of  $\lambda \cdot (s_{i,f}^{I,ms} + s_f^{P,ms} + s_i^{I,ws} + s_i^{P,ws})$ , which depends solely on  $COV$  thanks to A1 ii) and A3, and is denoted next by  $G_f(\cdot|COV)$ , for  $f = \tilde{f}$  and for any vector of covariates  $COV$ .

Similarly, from the sub-sample without an incumbent, the joint knowledge of  $H_f^{NI}(\cdot|COV)$  (for each  $f$ ) and  $\psi$  allows to identify  $\lambda \cdot \bar{V}_{i,f} + \log[\mathbb{E}[e^{\lambda \cdot (s_i^{I,ws} + s_i^{P,ws})}|COV_i = COV]]$  (for the workers  $i$  such that  $COV_i = COV$  and  $f_i^{inc} \notin S_i$ )<sup>55</sup> and the distribution of  $\lambda \cdot (s_{i,f}^{I,ms} + s_f^{P,ms})$  (for each  $f$ ), i.e. the distribution  $G_f^{ms}(\cdot|COV)$ . We stress that A3 guarantees that the distribution of  $\lambda \cdot (s_{i,f}^{I,ms} + s_f^{P,ms})$  depends solely on  $COV$  (as reflected in our notation).

From A1 ii), the worker-specific signals are drawn independently of the matching-specific signals: a standard deconvolution argument now implies that  $G^{ws}(\cdot|COV)$  is identified from  $G_f^{ms}(\cdot|COV)$  and  $G_f(\cdot|COV)$  for a given firm  $f$ . The latter distributions are identified at least for  $f = \tilde{f}$ , which allows thus to identify  $G^{ws}(\cdot|COV)$ .

Formally, the deconvolution argument works as follows:<sup>56</sup> For a given CDF  $H$ , let  $\widehat{H}$  denote the Fourier transform of  $H$ . We have then a bijection between  $H$  and  $\widehat{H}$  characterized by the two relations  $\widehat{H}(\xi) = \int_{-\infty}^{+\infty} H(x)e^{-i\xi x} dx$  and  $H(x) = (\int_{-\infty}^{+\infty} \widehat{H}(\xi)e^{+i\xi x} d\xi) / \sqrt{2\pi}$ . 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. A1 ii) and A3 imply then that  $\widehat{G}_f(\xi|COV) = \widehat{G}_f^{ms}(\xi|COV) \cdot \widehat{G}^{ws}(\xi|COV)$ . Finally  $G^{ws}(\cdot|COV)$  is formally characterized as a function of the CDF  $G_f(\cdot|COV)$  and  $G_f^{ms}(\cdot|COV)$  by

<sup>54</sup>Given the structure of firms' payoff function, the scalar  $\lambda \cdot \bar{V}_{i,\tilde{f}}$  depends on  $i$  only through  $COV_i$  and the fact that  $f_i^{inc} = \tilde{f}$ .

<sup>55</sup>Again, in such sub-sample, the scalar  $\lambda \cdot \bar{V}_{i,f}$  depends on  $i$  only through  $COV_i$  and the fact that  $f_i^{inc} \neq f$ .

<sup>56</sup>In the structural econometrics of auctions, Li et al. (2000) are the first to develop a nonparametric estimation procedure based on Fourier transformations.

$$G^{ws}(x|COV) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{\int_{-\infty}^{+\infty} G_f(x'|COV) e^{-i\xi x'} dx'}{\int_{-\infty}^{+\infty} G_f^{ms}(x''|COV) e^{-i\xi x''} dx''} \cdot e^{+i\xi x} d\xi.$$

From  $G^{ws}(\cdot|COV)$ , we identify  $\log[\mathbb{E}[e^{\lambda \cdot (s_i^{I,ws} + s_i^{P,ws})} | COV_i = COV]]$  and then  $\lambda \cdot \bar{V}_{i,f}$  (given  $COV_i = COV$  and  $f \neq f_i^{inc}$ ).

So far, we have established that the function  $\psi$ , the distributions  $G^{ws}(\cdot|COV)$  and  $G_f^{ms}(\cdot|COV)$ ,  $f = 1, \dots, F$  (for any covariate  $COV \in \mathcal{S}_{COV}$ ), the scalars  $\lambda \cdot \bar{V}_{i,f}$ , for any pair  $(i, f)$  such that  $COV_i \in \mathcal{S}_{COV}$  and  $f_i^{inc} \neq f$  and also for the pairs  $(i, \tilde{f})$  such that  $COV_i \in \mathcal{S}_{COV}$  and  $f_i^{inc} = f$ .

What remains to be checked is that we are able to identify  $\lambda \cdot \bar{V}_{i,f}$  for the pairs  $(i, f)$  with  $f \neq \tilde{f}$  such that  $COV_i \in \mathcal{S}_{COV}$  and  $f_i^{inc} = f$ .

It is now that we use TA5 and the fact that our payoff structure in such that  $\bar{V}_{i,f}$  can be expressed as  $\beta_f + (\ln(E_{\epsilon_{i,f}}[e^{\lambda \cdot \epsilon_{i,f}}]))/\lambda$  plus a term that depends on the pair  $(i, f)$  only through the covariates  $(\mathbf{x}_{i,f}, \mathbf{z}_{i,f})$  and the incumbency indicator  $INC_{i,f}$ .

Consider a pair  $(i, f)$  with  $f \neq \tilde{f}$  such that  $COV_i = (\mathbf{x}_{i,f'}, \mathbf{z}_{i,f'})_{f'=1, \dots, F} \in \mathcal{S}_{COV}$  and  $f_i^{inc} = f$ . From TA5, there exists  $COV' = (\mathbf{x}_{f'}, \mathbf{z}_{f'})_{f'=1, \dots, F} \in \mathcal{S}_{COV}$  such that  $(\mathbf{x}_{\tilde{f}}, \mathbf{z}_{\tilde{f}}) = (\mathbf{x}_{i,f}, \mathbf{z}_{i,f})$ . Furthermore, since  $COV' \in \mathcal{S}_{COV}$  and given  $\tilde{f}$  in TA4 ii), we have shown previously that for any  $\tilde{i}$  such that  $COV_{\tilde{i}} = COV'$  and  $f_{\tilde{i}}^{inc} = \tilde{f}$ , then  $\bar{V}_{\tilde{i}, \tilde{f}}$  is identified.

Consider then a pair  $(i^*, f)$  with  $COV_{i^*} = COV_i \in \mathcal{S}_{COV}$  and  $f_{i^*}^{inc} \neq f$  and a pair  $(\tilde{i}^*, \tilde{f})$  with  $COV_{\tilde{i}^*} = COV' \in \mathcal{S}_{COV}$  and  $f_{\tilde{i}^*}^{inc} \neq \tilde{f}$ . We have already shown that  $\bar{V}_{i^*, f}$  and  $\bar{V}_{\tilde{i}^*, \tilde{f}}$  are identified. From our payoff structure in (3), we have then<sup>57</sup>

$$\bar{V}_{i,f} - \bar{V}_{\tilde{i}, \tilde{f}} = \bar{V}_{i^*, f} - \bar{V}_{\tilde{i}^*, \tilde{f}} \quad (\text{B.1})$$

and  $\bar{V}_{i,f}$  is thus identified which concludes the proof of Proposition 4.1.

**Remark:** The equality (B.1) remains valid if the vector  $\beta_x$  is firm-specific.<sup>58</sup> However, it will not be satisfied if the parameter  $\beta_{inc}$  is firm-specific. Fundamentally, if we only assume that there is a single firm that is an incumbent (as it is assumed in TA4), then we have to assume that the incumbency indicator  $INC_{i,f}$  does not interact with the variable  $f$ .

**Assumption A6 under the parametrization**  $\tau(w) = \tau \cdot \log(w)$ :

We consider below the linear-log specification for the incentive effect  $\tau(w) = \tau \cdot \log(w)$  and let  $\theta = \lambda \cdot \tau$ .

Take a given covariate  $COV \in \mathcal{S}_{COV}$  and let us assume that the following CDFs are known:  $H_f^{NI}(\cdot|COV)$  (for each  $f$ ),  $H_{\tilde{f}}^{inc}(\cdot|COV, \tilde{f})$  and  $H_{\tilde{f}'}^{ACT}(\cdot|COV, \tilde{f}')$ . Assume A1-A5 and TA1-TA3. Then we show that the parameter  $\theta$  or equivalently the function  $\psi$  is identified.

Let  $\{\bar{G}_f^{ms}(\cdot|COV)\}_{f=1, \dots, F}$  and  $\bar{G}^{ws}(\cdot|COV)$  denote the distribution of  $\lambda(s_{i,f}^{I,ms} + s_{i,f}^{P,ms})/(1-\theta)$  and  $\lambda(s_i^{I,ws} + s_i^{P,ws})/(1-\theta)$  conditional on  $COV_i = COV$ , respectively. Let  $\bar{G}^{ws}(\cdot|COV, f, \tilde{x})$  denote the distribution of  $\lambda(s_{i,f}^{I,ws} + s_{i,f}^{P,ws})/(1-\theta)$  conditional on  $\lambda(s_{i,f}^{I,ms} + s_{i,f}^{P,ms} + s_i^{I,ws} + s_i^{P,ws}) =$

<sup>57</sup>From (3), we get that both terms in (B.1) are equal to  $\beta_f - \beta_{\tilde{f}}$ .

<sup>58</sup>If we use the natural corresponding notation  $\beta_x^f$  and  $\beta_x^{\tilde{f}}$ , then we get from (3) that both terms in (B.1) are equal to  $\beta_f - \beta_{\tilde{f}} + (\beta_x^f - \beta_x^{\tilde{f}}) \cdot \mathbf{x}_{i,f}$ .

$(1 - \theta) \cdot \tilde{x}$ . For any  $\tilde{x} \in \mathbb{R}$ , let

$$U_{\tilde{f}}(\tilde{x}|COV) := \left[ \log \left[ \int_{-\infty}^{+\infty} e^{(1-\theta) \cdot x} d\bar{G}^{ws}(x|COV, f, \tilde{x}) \right] \right] / (1 - \theta). \quad (\text{B.2})$$

**Remark:** Note that we consider here the product of  $\lambda/(1 - \theta)$  times the aggregate signals  $\bar{s}_{i,f}$ . This contrasts with our proof under the general case (and with A6) where we have considered the distribution of the product of  $\lambda$  and the aggregate signals  $\bar{s}_{i,f}$ .

In an auction for a given worker  $i$  without any incumbent and with  $COV_i = COV$ , the cutoff aggregate signal that makes firm  $f$  bid exactly until  $w$ , denoted next by  $\hat{s}_{i,f}^{[NI]}(w)$ , is characterized by

$$\frac{\lambda}{1 - \theta} \cdot \hat{s}_{i,f}^{[NI]}(w) = \log(w) - \frac{1}{1 - \theta} \cdot \log \left( \int_{-\infty}^{+\infty} e^{(1-\theta) \cdot x} \cdot d\bar{G}^{ws}(x|COV) \right) - \frac{\lambda}{1 - \theta} \cdot \bar{V}_{i,f}.$$

From the CDF  $H_f^{NI}(\cdot|COV)$  and given the normalization A1 iii), then we identify the distributions  $\{\bar{G}_f^{ms}(\cdot|COV)\}_{f=1,\dots,F}$  and the constants  $\left[ \log \left( \int_{-\infty}^{+\infty} e^{(1-\theta) \cdot x} \cdot d\bar{G}^{ws}(x|COV) \right) + \lambda \cdot \bar{V}_{i,f} \right] / (1 - \theta)$  for each  $f$ .

In an auction for a given worker  $i$  with  $COV_i = COV$  and with the set of participants  $S_i = \{\tilde{f}, \tilde{f}'\}$  where  $f_i^{inc} = \tilde{f}$ , the cutoff aggregate signal that makes the incumbent firm  $\tilde{f}$  bid exactly until  $w$ , denoted next by  $\hat{s}_{i,\tilde{f}}^{inc}(w)$ , is characterized by

$$\frac{\lambda}{(1 - \theta)} \cdot \hat{s}_{i,\tilde{f}}^{inc}(w) = \log(w) - \frac{\lambda}{(1 - \theta)} \cdot \bar{V}_{i,\tilde{f}}$$

and the cutoff aggregate signal that makes the non-incumbent firm  $\tilde{f}'$  bid exactly until  $w$ , denoted next by  $\hat{s}_{i,\tilde{f}'}^{[ACT]}(w)$ , is characterized by

$$\frac{\lambda}{1 - \theta} \cdot \hat{s}_{i,\tilde{f}'}^{[ACT]}(w) = \log(w) - U_{\tilde{f}}(\log(w) - \frac{\lambda}{1 - \theta} \cdot \bar{V}_{i,\tilde{f}}|COV) - \frac{\lambda}{1 - \theta} \cdot \bar{V}_{i,\tilde{f}'} \quad (\text{B.3})$$

where the function  $U_f(\cdot|COV)$  has been defined in (B.2).

From the CDF  $H_{\tilde{f}}^{inc}(\cdot|COV, \tilde{f})$  and given the normalization A1 iii), then we obtain that the scalar  $\lambda \bar{V}_{i,\tilde{f}} / (1 - \theta)$  and the distribution of  $\lambda [s_{i,\tilde{f}}^{I,ms} + s_{i,\tilde{f}}^{I,ms} + s_{i,\tilde{f}}^{I,ms} + s_{i,\tilde{f}}^{I,ms}] / (1 - \theta)$  are identified and then from the same deconvolution argument as in the proof for the general case, we identify the CDFs  $\bar{G}^{ws}(\cdot|COV)$  and  $\bar{G}^{ws}(\cdot|COV, f, \tilde{x})$  for each  $\tilde{x} \in \mathbb{R}$ .

In equation (B.3), we have shown that the scalars  $\lambda \bar{V}_{i,\tilde{f}'} / (1 - \theta)$  and  $\lambda \bar{V}_{i,\tilde{f}} / (1 - \theta)$  are identified. We have also shown that the distribution of the left-hand term  $\lambda \bar{s}_{i,\tilde{f}'} / (1 - \theta)$  is identified since it corresponds to  $\bar{G}_{\tilde{f}'}^{ms}(\cdot|COV)$  thanks to the exclusion restriction A3 that guarantees that for a given set of covariates  $COV_i$ , the distribution of  $s_{i,f}^{I,ms} + s_{i,f}^{P,ms}$  does not depend on  $f_i^{inc}$ .

From the equilibrium equation (B.3), the equilibrium drop-out price of the non-incumbent  $\tilde{f}'$  facing only the incumbent  $\tilde{f}$  as a function of  $\lambda \bar{s}_{i,\tilde{f}'} / (1 - \theta)$  can thus be expressed as  $\Lambda(\lambda \bar{s}_{i,\tilde{f}'} / (1 - \theta), \theta)$ , where the function  $\Lambda$  has been identified from our previous identification steps.<sup>59</sup> We then use the following result (see Proposition 2 in Gollier (2001)):

<sup>59</sup>Formally the function  $\Lambda(\cdot, \theta)$  which is defined by (B.3) depends on the scalars  $\frac{\lambda}{1 - \theta} \cdot \bar{V}_{i,f}$  for  $f = \tilde{f}', \tilde{f}$  and on

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

Take a given quantile of the distribution  $H_{\tilde{f}'}^{ACT}(\cdot | \text{COV}, \tilde{f})$  (which has been shown to be identified). Let  $w^*$  denote the cutoff wage associated to this quantile.

Let us now consider the corresponding quantile of the distribution  $\overline{G}_{\tilde{f}'}^{ms}(\cdot | \text{COV})$  (which has been shown to be identified), denoted next by  $bs^*$ . The non-incumbent  $\tilde{f}'$  facing the incumbent  $\tilde{f}$  and such that  $\lambda \bar{s}_{i, \tilde{f}'} = (1 - \theta)bs^*$  should bid  $\Lambda(bs^*, \theta)$

Given the expression of  $U_{\tilde{f}}(\cdot | \text{COV})$ , the fact that  $\overline{G}^{ws}(\cdot | \text{COV}, f, \tilde{x})$  has a positive variance for any  $\tilde{x} \in \mathbb{R}$  (thanks to our full support assumption TA2 and the independence assumption A1 ii) and Lemma 1, we get that  $\Lambda(bs^*, \theta)$  is increasing in  $\theta$ . There is thus at most one value of  $\theta$  such that  $\Lambda(bs^*, \theta) = w^*$  and  $\theta$  is thus identified.

## C Derivation of the augmented performance equation (6)

From the law of iterated expectations, we have  $\mathbb{E}[s_{i, f_i}^{P, ms} | \mathcal{I}_i] = \mathbb{E}[\mathbb{E}[s_{i, f_i}^{P, ms} | \mathcal{I}_i, s_i^{I, ms} + s_{i, f_i}^{P, ms}] | \mathcal{I}_i]$  and  $\mathbb{E}[s_{i, f_i}^{P, ws} | \mathcal{I}_i] = \mathbb{E}[\mathbb{E}[s_{i, f_i}^{P, ws} | \mathcal{I}_i, s_i^{I, ws} + bs_i^{P, ws}] | \mathcal{I}_i]$ . Next we use that the bidding strategy of each firm  $f$  for worker  $i$  depends on the vector of signals  $\{s_{i, f}\}_{i=1, \dots, N}$  only through the aggregate signal  $\bar{s}_{i, f}$  (Proposition A.1). Furthermore, given A1, A3 and A5, the pair of signals  $(s_{i, f}^{I, ms}, s_{i, f}^{P, ms})$  (resp.  $(s_i^{I, ws}, s_i^{P, ws})$ ) is drawn independently of the other signals and of  $f_i^{inc}$  and  $S_i$ : its distribution depends solely on  $\text{COV}_i$ . We obtain then that the distribution of  $s_{i, f_i}^{P, ms}$  conditional on  $(\mathcal{I}_i, s_{i, f_i}^{I, ms} + s_{i, f_i}^{P, ms})$  (resp.  $s_i^{P, ms}$  conditional on  $(\mathcal{I}_i, s_i^{I, ws} + s_i^{P, ws})$ ) is the same as the distribution of  $s_{i, f_i}^{P, ms}$  conditional on  $(\text{COV}_i, s_{i, f_i}^{I, ms} + s_{i, f_i}^{P, ms})$  (resp.  $s_i^{P, ws}$  conditional on  $(\text{COV}_i, s_i^{I, ws} + s_i^{P, ws})$ ). In particular, the auction outcome  $(w_i, f_i^w)$  does not influence this conditional distribution. We have thus shown that  $\mathbb{E}[s_{i, f_i}^{P, ms} | \mathcal{I}_i] = \mathbb{E}[\mathbb{E}[s_{i, f_i}^{P, ms} | \text{COV}_i, s_{i, f_i}^{I, ms} + s_{i, f_i}^{P, ms}] | \mathcal{I}_i]$  and  $\mathbb{E}[s_{i, f_i}^{P, ws} | \mathcal{I}_i] = \mathbb{E}[\mathbb{E}[s_{i, f_i}^{P, ws} | \text{COV}_i, s_i^{I, ws} + bs_i^{P, ws}] | \mathcal{I}_i]$

Given the expressions of  $\mathbb{E}[s_{i, f_i}^{P, ms} | \text{COV}_i, s_{i, f_i}^{I, ms} + s_{i, f_i}^{P, ms}]$  and  $\mathbb{E}[s_{i, f_i}^{P, ws} | \text{COV}_i, s_i^{I, ws} + bs_i^{P, ws}]$  in A7, the conditional expectation  $\mathbb{E}[u_{i, f_i^w} | \mathcal{I}_i] = \mathbb{E}[s_{i, f_i^w}^{P, ms} | \mathcal{I}_i] + \mathbb{E}[s_{i, f_i^w}^{P, ws} | \mathcal{I}_i]$  now takes the following form

$$\mathbb{E}[u_{i, f_i^w} | \mathcal{I}_i] = \sum_{l=1}^L \left[ \gamma_{l, f_i^w}^{ms}(\text{COV}_i) \cdot CF_l^{ms}[l] + \gamma_l^{ws}(\text{COV}_i) \cdot CF_l^{ws}[l] \right]. \quad (\text{C.1})$$

Combined with the performance equation (1) and the definition of the control terms in the augmented performance equation (6), we obtain thus that the error term  $\xi_{i, f_i^w}$  is equal to  $u_{i, f_i^w} - \mathbb{E}[u_{i, f_i^w} | \mathcal{I}_i]$ .

Let us now show that the control terms  $CF_l^{ms}[l]$  and  $CF_l^{ws}[l]$ ,  $l = 1, \dots, L$ , are identified from the auction data. Thanks to Proposition 4.1 and Corollary 1 (in Appendix A), we have already identified from the bidding data the distribution of the bidders' aggregate signals  $\lambda \cdot \bar{s}_{i, f}$ ,<sup>60</sup> and the equilibrium strategy of each firm  $f \in S_i$  in the auction for worker  $i$  as a function of its signal  $\lambda \cdot \bar{s}_{i, f}$ . Conditional on  $\text{COV}_i$ ,  $f_i^{inc}$  and  $S_i$ , the auction outcome  $(w_i, f_i^w)$  is thus fully

the distribution  $\overline{G}^{ws}(\cdot | \text{COV}, \tilde{f}, \tilde{x})$  for any  $\tilde{x} \in \mathbb{R}$ . Given those identified primitives, the dependence on  $\theta$  of the right-hand term is only through the function  $U_f$ .

<sup>60</sup>If the auction  $i$  involves an incumbent, the distribution of  $\lambda \cdot \bar{s}_{i, f_i^{inc}}$  is identified as it is the sum of two independent variables,  $\lambda \cdot (s_{i, f_i^{inc}}^{I, ms} + s_{i, f_i^{inc}}^{P, ms})$  and  $\lambda \cdot (s_i^{I, ws} + s_i^{P, ws})$ , whose distributions, that is to say,  $G_{f_i^{inc}}^{ms}(\cdot | \text{COV}_i)$  and  $G^{ws}(\cdot | \text{COV}_i)$  are identified from Proposition 4.1.

determined as a function of the vector  $\lambda \cdot \bar{s}_{i,f}$  for  $f = 1, \dots, F$ . Formally, we thus identify the joint distribution of  $\{\lambda \cdot \bar{s}_{i,f}\}_{f=1, \dots, F}$  and  $(w_i, f_i^w)$ , conditional on  $COV_i, f_i^{inc}$  and  $S_i$ . Bayesian updating allows in turn to determine the distribution of  $\lambda \cdot (s_{i,f_i^w}^{I,ms} + s_{i,f_i^w}^{P,ms})$  and  $\lambda \cdot (s_i^{I,ws} + s_i^{P,ws})$  conditional on  $\mathcal{A}_i$ . From the various moments of those conditional distributions, we identify the control function terms  $CF_i^k[l]$ , for  $k = ms, ws$  and  $l = 1, \dots, L$ . Note that if the auction involves a reservation wage, we still identify the distribution of the auction outcome and a similar Bayesian updating exercise can be done to identify the control terms.

We conclude by noting that there is no correlation between the error term and the explanatory variables in (6) insofar as the set  $\mathcal{A}_i$  includes all explanatory variables: either those appearing directly in the original performance equation (i.e., the variables  $f_i^w, \mathbf{x}_{i,f_i^w}, INC_{i,f_i^w}$  and  $w_i$ ), or those appearing indirectly through the control terms  $CF_i^k[l]$ ,  $k = ms, ws, l = 1, \dots, L$  that have been shown to be known functions of  $\mathcal{A}_i$ .

# Supplementary Appendix (for online publication only)

## D Assumption A7 under the Gaussian structure

If the vector  $(s_f^{I,ms}, s_f^{P,ms})$  follows a bivariate normal distribution with the variances  $\sigma_{I,ms}$  and  $\sigma_{P,ms}$  and the correlation coefficient  $\rho_{ms}$ , then we have  $\sigma_{ms} = \sqrt{\sigma_{I,ms}^2 + \sigma_{P,ms}^2 + 2 \cdot \rho_{ms} \sigma_{I,ms} \sigma_{P,ms}}$  and  $\mathbb{E}[s_{i,f}^{P,ms} | s_{i,f}^{I,ms} + s_{i,f}^{P,ms} = u] = \gamma^{ms} \cdot u$  with  $\gamma^{ms} = \frac{[\sigma_{P,ms}]^2 + \rho_{ms} \sigma_{I,ms} \sigma_{P,ms}}{\sigma_{ms}^2}$ . Note that  $\gamma^{ms} \in [0, 1]$  if  $\rho_{ms} \geq 0$ .<sup>61</sup> Analog results hold if  $(s_f^{I,ws}, s_f^{P,ws})$  follow a bivariate normal distribution. On the whole, if both  $(s_f^{I,ms}, s_f^{P,ms})$  and  $(s_f^{I,ws}, s_f^{P,ws})$  follow bivariate normal distributions with positive correlation coefficients, then A7 hold automatically and with  $L = 1$ .

## E Details on the computation of the control functions

In this section, we detail how the control terms  $CF_i^k[l]$ ,  $l = 1, \dots, L$ ,  $k = ms, ws$ , are characterized when the bidding history of the incumbent is observed by the econometrician. To obtain the precise expressions of these functions, we need to perform Bayesian updating conditional on the information set  $\mathcal{I}_i$  given the equilibrium behavior that is characterized in Proposition A.1. We also illustrate our computations under the parametric specification introduced in Section 4.3. Here, to get explicit and tractable expressions, we use a series of well known properties on (truncated) normally distributed variables (which can be found e.g. in 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 its bidding history. Following Appendix A's nomenclature, we distinguish thus four cases: A) There is no incumbent; B) The winner is not the incumbent and the incumbent did not enter the auction (at the reserve  $W_i^r$ ); C) The winner is not the incumbent and the incumbent has dropped out at the wage  $w_i^{inc} \in [W_i^r, w_i]$ ; D) The winner is the incumbent.

Below we then use some notation introduced in our equilibrium analysis in Appendix A. Note that we restore the index  $i$  in the corresponding equilibrium cut-off strategies (as the function of the bidder's aggregate signal):  $w_{i,f}(\cdot)$ ,  $w_{i,f}[NI](\cdot)$ ,  $w_{i,f}[NE](\cdot)$ ,  $w_{i,f}[DROP](\cdot)$ ;  $w'$  and  $w_{i,f}[ACT]$ .

**Case A:** In the absence of an incumbent, worker  $i$  is sold to firm  $f_i^w$  at price  $w_i > W_i^r$  (resp.  $w_i = W_i^r$ ) if and only if  $w_{i,f_i^w}[NI](\bar{s}_{i,f_i^w}) \geq w_i$  and  $\max_{f \in \mathcal{S}_i \setminus \{f_i^w\}} \{w_{i,f}[NI](\bar{s}_{i,f})\} = w_i$  (resp.  $\max_{f \in \mathcal{S}_i \setminus \{f_i^w\}} w_{i,f}[NI](\bar{s}_{i,f}) \leq W_i^r$ ). Given the independence assumptions in A1 and A5, we have that conditional on  $\mathcal{I}_i$ ,  $\lambda \cdot \bar{s}_{i,f_i^w}$  is distributed according to the distribution  $G_{f_i^w}^{ms}(\cdot | COV_i)$  truncated from below at the signal  $[\lambda \cdot w_{i,f_i^w}[NI]]^{-1}(w_i)$ , i.e., it has the distribution function  $\frac{G_{f_i^w}^{ms}(\cdot | COV_i) - G_{f_i^w}^{ms}(\lambda \cdot [w_{i,f_i^w}[NI]]^{-1}(w_i) | COV_i)}{1 - G_{f_i^w}^{ms}(\lambda \cdot [w_{i,f_i^w}[NI]]^{-1}(w_i) | COV_i)}$  on the interval  $[\lambda \cdot [w_{i,f_i^w}[NI]]^{-1}(w_i), +\infty)$ . Since there is no incumbent in this case, our independence assumptions guarantee that there is no updating on  $s_i^{I,ws} + s_i^{P,ws}$  and the terms  $CF_i^{ws}[l]$ ,  $l = 1, \dots, L$ , depends thus on  $\mathcal{I}_i$  only through  $COV_i$ .

<sup>61</sup>If  $\rho_{ms} < 0$ , then we may have  $\gamma^{ms} \notin [0, 1]$ . The inequality  $\gamma^{ms} < 0$  holds if and only if  $\rho_{ms} < -\frac{\sigma_{P,ms}}{\sigma_{I,ms}}$ , which can never occur if  $\sigma_{P,ms} \geq \sigma_{I,ms}$ . The inequality  $\gamma^{ms} > 1$  holds if and only if  $\rho_{ms} < -\frac{\sigma_{I,ms}}{\sigma_{P,ms}}$ , which can never occur if  $\sigma_{I,ms} \geq \sigma_{P,ms}$ .

More precisely, we have

$$CF_i^{ms}[L] = \int_{\lambda \cdot [w_{i,f_i^w}[NE]]^{-1}(w_i)}^{+\infty} u^L \cdot \frac{d[G_{f_i^w}^{ms}(u|COV_i)]}{1 - G_{f_i^w}^{ms}(\lambda \cdot [w_{i,f_i^w}[NE]]^{-1}(w_i)|COV_i)} \quad \text{and} \quad CF_i^{ws}[L] = \int_{-\infty}^{+\infty} u^L \cdot d[G^{ws}(u|COV_i)]. \quad (\text{E.1})$$

Under our parametric specification, and noting that  $\log(\int_{-\infty}^{+\infty} e^u d\Phi(\frac{u}{\sigma_{ws}})) = \frac{\sigma_{ws}^2}{2}$ , we obtain the expression in (8).

The inverse Mills ratio in (9) is obtained from the well known expression of the mean of a truncated normal distribution.

**Case B:** From Proposition A.1 we know that the incumbent  $f_i^{inc}$  does not enter the auction for worker  $i$  and that worker  $i$  is sold to firm  $f_i^w$  at price  $w_i > W_i^r$  (resp.  $w_i = W_i^r$ ) if and only if  $w_{i,f_i^{inc}}(\bar{s}_{i,f_i^{inc}}) \equiv w_{i,f_i^{inc}}^* < W_i^r$ ,  $w_{i,f_i^w}[NE](\bar{s}_{i,f_i^w}) \geq w_i$  and  $\max_{f \in S_i \setminus \{f_i^w, f_i^{inc}\}} \{w_{i,f}[NE](\bar{s}_{i,f})\} = w_i$  (resp.  $\max_{f \in S_i \setminus \{f_i^w, f_i^{inc}\}} \{w_{i,f}[NE](\bar{s}_{i,f})\} < W_i^r$ ). Given the independence assumptions in A1 and A5, we have on the one hand that conditional on  $\mathcal{S}_i$ ,  $\lambda \cdot \bar{s}_{i,f_i^w}$  is distributed according to the distribution  $G_{f_i^w}^{ms}(\cdot|COV_i)$  truncated from below at the signal realization  $\lambda \cdot [w_{i,f_i^w}[NE]]^{-1}(w_i)$ ,

i.e., it has the distribution function  $\frac{G_{f_i^w}^{ms}(\cdot|COV_i) - G_{f_i^w}^{ms}(\lambda \cdot [w_{i,f_i^w}[NE]]^{-1}(w_i)|COV_i)}{1 - G_{f_i^w}^{ms}(\lambda \cdot [w_{i,f_i^w}[NE]]^{-1}(w_i)|COV_i)}$  on the interval  $[\lambda \cdot [w_{i,f_i^w}[NE]]^{-1}(w_i), +\infty)$ . On the other hand, we have that the distribution of  $\lambda \cdot \bar{s}_{i,f_i^{inc}}$  conditional on  $\mathcal{S}_i$  is distributed according to the distribution  $G_{f_i^{inc}}(\cdot|COV_i)$  truncated from above at the signal  $\lambda \cdot w_{i,f_i^{inc}}^{-1}(W_i^r)$ , where the CDF  $G_{f_i^{inc}}(\cdot|COV_i)$  denotes the distribution of  $\lambda \cdot \bar{s}_{i,f_i^{inc}}$  which depends on  $i$  only through  $f_i^{inc}$  and  $COV_i$ . We have then

$$CF_i^{ms}[L] = \int_{\lambda \cdot [w_{i,f_i^w}[NE]]^{-1}(w_i)}^{+\infty} u^L \cdot \frac{d[G_{f_i^w}^{ms}(u|COV_i)]}{1 - G_{f_i^w}^{ms}(\lambda \cdot [w_{i,f_i^w}[NE]]^{-1}(w_i)|COV_i)} \quad (\text{E.2})$$

and

$$\begin{aligned} CF_i^{ws}[L] &= \int_{-\infty}^{\lambda \cdot w_{i,f_i^{inc}}^{-1}(W_i^r)} \mathbb{E}[(\lambda \cdot (s_i^{I,ws} + s_i^{P,ws}))^L | \lambda \cdot \bar{s}_{i,f_i^{inc}} = \tilde{u}, COV_i] \cdot \frac{dG_{f_i^{inc}}(\tilde{u}|COV_i)}{G_{f_i^{inc}}(\lambda \cdot w_{i,f_i^{inc}}^{-1}(W_i^r)|COV_i)} \\ &= \int_{-\infty}^{\lambda \cdot w_{i,f_i^{inc}}^{-1}(W_i^r)} \left[ \int_{-\infty}^{+\infty} u^L \cdot \frac{g^{ws}(u|COV_i) g_{f_i^{inc}}^{ms}(\tilde{u} - u|COV_i)}{G_{f_i^{inc}}(\lambda \cdot w_{i,f_i^{inc}}^{-1}(W_i^r)|COV_i)} du \right] \cdot d\tilde{u}. \end{aligned}$$

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

$$CF_i^{ms}[1] = \sigma_{ms} \cdot \frac{\phi\left(\frac{\lambda \cdot [w_{i,f_i^w}[NE]]^{-1}(w_i)}{\sigma_{ms}}\right)}{1 - \Phi\left(\frac{\lambda \cdot [w_{i,f_i^w}[NE]]^{-1}(w_i)}{\sigma_{ms}}\right)} \quad \text{and} \quad CF_i^{ws}[1] = -\frac{\sigma_{ws}^2}{\sqrt{\sigma_{ms}^2 + \sigma_{ws}^2}} \cdot \frac{\phi\left(\frac{\lambda \cdot w_{i,f_i^{inc}}^{-1}(W_i^r)}{\sqrt{\sigma_{ms}^2 + \sigma_{ws}^2}}\right)}{\Phi\left(\frac{\lambda \cdot w_{i,f_i^{inc}}^{-1}(W_i^r)}{\sqrt{\sigma_{ms}^2 + \sigma_{ws}^2}}\right)}. \quad (\text{E.3})$$

To pursue the computation, let us derive the expressions of  $\lambda \cdot [w_{i,f_i^w}[NE]]^{-1}(w_i)$  and  $\lambda \cdot w_{i,f_i^{inc}}^{-1}(W_i^r)$ , whose tractable expressions can then be plugged into (E.3).

First we note that  $\lambda \cdot w_{i,f_i^{inc}}^{-1}(W_i^r) = (1 - \theta) \log(W_i^r) - \lambda \cdot \bar{V}_{i,f_i^{inc}}$ . Second, we have

$$\lambda \cdot [w_{i,f_i^w}[NE]]^{-1}(w_i) = (1 - \theta) \log(w_i) - \lambda \cdot \bar{V}_{i,f_i^w} - \log(\mathbb{E}[e^{\lambda \cdot [s_i^{I,ws} + s_i^{P,ws}]} | \lambda \cdot \bar{s}_{i,f_i^{inc}} < \lambda \cdot w_{i,f_i^{inc}}^{-1}(W_i^r)]). \quad (\text{E.4})$$



The expectation appearing in (E.4) is of the form  $\mathbb{E}[e^{y_1}|y_2 \leq a]$  with  $(y_1, y_2)$  following a bivariate centered normal distribution with the corresponding standard deviations  $\sigma_1 = \sigma_{ws}$  and  $\sigma_2 = \sqrt{\sigma_{ms}^2 + \sigma_{ws}^2}$  and the coefficient of correlation  $\rho = \frac{\sigma_{ws}}{\sqrt{\sigma_{ms}^2 + \sigma_{ws}^2}}$ , and with  $a = (1 - \theta) \log(W_i^r) - \lambda \cdot \bar{V}_{i,f_i^{inc}}$ . 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  $\mathbb{E}[e^{y_1}|y_2 = a] = \exp(\rho \cdot \frac{\sigma_1}{\sigma_2} \cdot a + (1 - \rho^2) \cdot \frac{\sigma_1^2}{2})$  and then that  $\mathbb{E}[e^{y_1}|y_2 \leq a] = \mathbb{E}[\exp(\rho \cdot \frac{\sigma_1}{\sigma_2} y_2)|y_2 \leq a] \cdot \exp((1 - \rho^2) \cdot \frac{\sigma_1^2}{2})$ . 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

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

Replacing  $\sigma_1$ ,  $\sigma_2$ ,  $\rho$  and  $a$  by the aforementioned values, (E.4) becomes

$$\lambda \cdot [w_{i,f_i^w}[NE]]^{-1}(w_i) = (1 - \theta) \log(w_i) - \lambda \cdot \bar{V}_{i,f_i^w} - \log\left(\frac{\Phi\left(\frac{[-\sigma_{ws}^2 + (1 - \theta) \log(W_i^r) - \lambda \cdot \bar{V}_{i,f_i^{inc}}]}{\sqrt{\sigma_{ms}^2 + \sigma_{ws}^2}}\right)}{\Phi\left(\frac{(1 - \theta) \log(W_i^r) - \lambda \cdot \bar{V}_{i,f_i^{inc}}}{\sqrt{\sigma_{ms}^2 + \sigma_{ws}^2}}\right)}\right) + \frac{\sigma_{ws}^2}{2} \quad (\text{E.5})$$

**Case C:** From Proposition A.1, the incumbent exit the auction at  $w_i^{inc} \geq W_i^r$  and the worker  $i$  is sold to firm  $f_i^w \neq f_i^{inc}$  at  $w_i > w_i^{inc}$  (resp.  $w_i = w_i^{inc}$ ) if and only if  $w_{i,f_i^{inc}}(\bar{s}_{i,f_i^{inc}}) = w_i^{inc}$ ,  $w_{i,f_i^w}[DRO]P](\bar{s}_{i,f_i^w}; w_i^{inc}) \geq w_i$  and  $\max_{f \in S_i \setminus \{f_i^w, f_i^{inc}\}} \{w_{i,f}[DRO]P](\bar{s}_{i,f}; w_i^{inc})\} = w_i$  (resp.  $\max_{f \in S_i \setminus \{f_i^w, f_i^{inc}\}} \{w_{i,f}[DRO]P](\bar{s}_{i,f}; w_i^{inc})\} \leq w_i$ ). Given our independence assumptions, we obtain that  $\lambda \cdot \bar{s}_{i,f_i^{inc}} = \lambda \cdot [w_{i,f_i^{inc}}]^{-1}(w_i^{inc})$  and that  $\lambda \cdot \bar{s}_{i,f_i^w}$  is distributed according to the distribution  $G_{f_i^w}^{ms}(\cdot|COV_i)$  truncated from below at the signal  $\lambda \cdot [w_{i,f}[DRO]P]^{-1}(w_i; w_i^{inc})$ .

In general, we obtain then the following expressions for the control terms:

$$CF_i^{ms}[l] = \int_{\lambda \cdot [w_{i,f}[DRO]P]^{-1}(w_i; w_i^{inc})}^{+\infty} u^l \cdot \frac{d[G_{f_i^w}^{ms}(u|COV_i)]}{1 - G_{f_i^w}^{ms}(\lambda \cdot [w_{i,f}[DRO]P]^{-1}(w_i; w_i^{inc})|COV_i)} \quad (\text{E.6})$$

and

$$CF_i^{ws}[l] = \int_{-\infty}^{+\infty} u^l \cdot \frac{g^{ws}(u|COV_i) g_{f_i^{inc}}^{ms}(\lambda \cdot [w_{i,f_i^{inc}}]^{-1}(w_i^{inc}) - u|COV_i)}{G_{f_i^{inc}}(\lambda \cdot w_{i,f_i^{inc}}^{-1}(W_i^r)|COV_i)} du \quad (\text{E.7})$$

Let us now detail the computations under our parametric structure.

We have  $\lambda \cdot [w_{i,f_i^{inc}}]^{-1}(w_i^{inc}) = (1 - \theta) \log(w_i^{inc}) - \lambda \cdot \bar{V}_{i,f_i^{inc}}$

From any  $f \neq f_i^{inc}$ , we have also from our equilibrium analysis:  $\lambda \cdot [w_{i,f}[DRO]P]^{-1}(w, w_i^{inc}) = (1 - \theta) \log(w) - \lambda \cdot \bar{V}_{i,f} - \log(\mathbb{E}[e^{\lambda \cdot [s_i^{I,ws} + s_i^{P,ws}]}] | \bar{s}_{f_i^{inc}} = w_{f_i^{inc}}^{-1}(w_i^{inc})])$

Note also that the term  $\mathbb{E}[e^{\lambda \cdot [s_i^{I,ws} + s_i^{P,ws}]} | \lambda \cdot \bar{s}_{f_i^{inc}} = \lambda \cdot w_{f_i^{inc}}^{-1}(w_i^{inc})]$  takes the form  $\mathbb{E}[e^{y_1}|y_2 = a]$  with  $(y_1, y_2)$  following a bivariate centered normal distribution with the corresponding standard deviations  $\sigma_1 = \sigma_{ws}$  and  $\sigma_2 = \sqrt{\sigma_{ms}^2 + \sigma_{ws}^2}$  and the coefficient of correlation  $\rho = \frac{\sigma_{ws}}{\sqrt{\sigma_{ms}^2 + \sigma_{ws}^2}}$  and with  $a = (1 - \theta) \log(w_i^{inc}) - \lambda \cdot \bar{V}_{i,f_i^{inc}}$ . Applying the computation presented in Case B, we obtain that

$$\log(\mathbb{E}[e^{\lambda \cdot [s_i^{I,ws} + s_i^{P,ws}]} | \bar{s}_{f,inc} = w_{f,inc}^{-1}(w_i^{inc})]) = \frac{\sigma_{ws}^2}{\sigma_{ms}^2 + \sigma_{ws}^2} \cdot [(1-\theta) \log(w_i^{inc}) - \lambda \cdot \bar{V}_{i,f,inc}] + \frac{\sigma_{ms}^2}{\sigma_{ms}^2 + \sigma_{ws}^2} \cdot \frac{\sigma_{ws}^2}{2}. \quad (\text{E.8})$$

Under our parametric specification, we obtain then

$$\lambda \cdot [w_{i,f}[DROP]]^{-1}(w; w_i^{inc}) = (1-\theta) \frac{\sigma_{ms}^2}{\sigma_{ms}^2 + \sigma_{ws}^2} \log(w) - \lambda \bar{V}_{i,f} + \frac{\sigma_{ws}^2}{\sigma_{ms}^2 + \sigma_{ws}^2} \cdot \lambda \cdot \bar{V}_{i,f,inc} - \frac{\sigma_{ms}^2}{\sigma_{ms}^2 + \sigma_{ws}^2} \cdot \frac{\sigma_{ws}^2}{2}.$$

Finally, we get the following expressions for the control terms

$$CF_i^{ms}[1] = \sigma_{ms} \cdot \frac{\phi(\lambda \cdot [w_{i,f_i^w}[DROP]]^{-1}(w_i; w_i^{inc}))}{1 - \Phi(\lambda \cdot [w_{i,f_i^w}[DROP]]^{-1}(w_i; w_i^{inc}))} \quad \text{and} \quad CF_i^{ws}[1] = \frac{\sigma_{ws}^2}{\sigma_{ms}^2 + \sigma_{ws}^2} \cdot [(1-\theta) \log(w_i^{inc}) - \lambda \cdot \bar{V}_{i,f_i^{inc}}].$$

The expression for  $CF_i^{ws}$  corresponds thus to (10).

**Case D:** From Proposition A.1, the incumbent wins the auction at  $w_i > W_i^r$  (resp.  $w_i = W_i^r$ ) if and only if  $w_{i,f,inc}(\bar{s}_{i,f,inc}) \geq w_i$  and  $\max_{f \in \mathcal{S}_i \setminus \{f_i^{inc}\}} \{w_{i,f}[ACT](\bar{s}_{i,f})\} = w_i$  (resp.  $\max_{f \in \mathcal{S}_i \setminus \{f_i^{inc}\}} \{w_{i,f}[ACT](\bar{s}_{i,f})\} \leq w_i$ ). Given our independence assumptions, we obtain that conditional on  $\mathcal{I}_i$ , the aggregate signal  $\lambda \cdot \bar{s}_{i,f,inc} = \lambda \cdot [w_{i,f,inc}]^{-1}(w_i^{inc})$  is distributed according to the distribution  $G_{f_i^{inc}}(\cdot | COV_i)$  truncated from below at the signal  $\lambda \cdot [w_{i,f,inc}]^{-1}(w_i^{inc})$ . Furthermore, conditional on a given realization for  $\lambda \cdot \bar{s}_{i,f,inc}$ , the distribution of the two independent aggregate signal  $\lambda \cdot (s_{i,f,inc}^{I,ms} + s_{i,f,inc}^{P,ms})$  and  $\lambda \cdot (s_i^{I,ws} + s_i^{P,ws})$  depend on  $\mathcal{I}_i$  only through  $COV_i$  and  $f_i^{inc}$ .

From the independence between worker and match-specific signals, we obtain by deconvolution:

$$\begin{aligned} CF_i^{ms}[l] &= \int_{-\infty}^{\lambda \cdot w_{i,f,inc}^{-1}(w_i)} \mathbb{E}[(\lambda \cdot (s_i^{I,ms} + s_i^{P,ms}))^l | \lambda \cdot \bar{s}_{i,f,inc} = \tilde{u}, COV_i] \cdot \frac{dG_{f_i^{inc}}(\tilde{u} | COV_i)}{G_{f_i^{inc}}(\lambda \cdot w_{i,f,inc}^{-1}(w_i) | COV_i)} \\ &= \int_{-\infty}^{\lambda \cdot w_{i,f,inc}^{-1}(W_i^r)} \left[ \int_{-\infty}^{+\infty} u^l \cdot \frac{g^{ms}(u | COV_i) g_{f_i^{inc}}^{ws}(\tilde{u} - u | COV_i)}{G_{f_i^{inc}}(\lambda \cdot w_{i,f,inc}^{-1}(w_i) | COV_i)} du \right] \cdot d\tilde{u} \end{aligned}$$

and

$$\begin{aligned} CF_i^{ws}[l] &= \int_{-\infty}^{\lambda \cdot w_{i,f,inc}^{-1}(w_i)} \mathbb{E}[(\lambda \cdot (s_i^{I,ws} + s_i^{P,ws}))^l | \lambda \cdot \bar{s}_{i,f,inc} = \tilde{u}, COV_i] \cdot \frac{dG_{f_i^{inc}}(\tilde{u} | COV_i)}{G_{f_i^{inc}}(\lambda \cdot w_{i,f,inc}^{-1}(W_i^r) | COV_i)} \\ &= \int_{-\infty}^{\lambda \cdot w_{i,f,inc}^{-1}(w_i)} \left[ \int_{-\infty}^{+\infty} u^l \cdot \frac{g^{ws}(u | COV_i) g_{f_i^{inc}}^{ms}(\tilde{u} - u | COV_i)}{G_{f_i^{inc}}(\lambda \cdot w_{i,f,inc}^{-1}(w_i) | COV_i)} du \right] \cdot d\tilde{u}. \end{aligned}$$

Under our parametric specification, the expression of the cutoff aggregate signal that makes the incumbent firm bid exactly until  $w$ , denoted next by  $\hat{s}_i^{inc}(w)$ , is given by

$$\lambda \cdot \hat{s}_i^{inc}(w) = (1-\theta) \log(w) - \lambda \cdot \bar{V}_{i,f,inc}. \quad (\text{E.9})$$

Under this parametric specification,, We can infer that conditional on  $\mathcal{I}_i$ , the aggregate signal of the winning firm,  $\lambda \cdot \bar{s}_{i,f,inc}$ , is distributed according to a centered normal distribution

(with variance  $\sigma_{ms}^2 + \sigma_{ws}^2$ )<sup>62</sup> truncated from below at  $\lambda \cdot \widehat{s}_i^{inc}(w_i)$ . Furthermore, the distribution of  $\lambda \cdot (s_{i,f}^{I,ms} + s_{i,f}^{P,ms})$  (resp.  $\lambda \cdot (s_i^{I,ws} + s_i^{P,ws})$ ) conditional on  $\mathcal{A}_i$  and  $\lambda \cdot \bar{s}_{i,f}^{inc} = u$  is a normal distribution with mean  $\sigma_{ms}^2 \cdot u / (\sigma_{ms}^2 + \sigma_{ws}^2)$  (resp.  $\sigma_{ws}^2 \cdot u / (\sigma_{ms}^2 + \sigma_{ws}^2)$ ). Combining these two facts gives the following control function terms:  $CF_i^{ws} = \sigma_{ws}^2 \cdot CF_i^{ms} / \sigma_{ms}^2$  and

$$CF_i^{ms} = \int_{\lambda \cdot \widehat{s}_i^{inc}(w_i)}^{+\infty} \frac{\sigma_{ms}^2}{\sigma_{ms}^2 + \sigma_{ws}^2} \cdot u \cdot \frac{d\Phi\left(\frac{u}{\sqrt{\sigma_{ms}^2 + \sigma_{ws}^2}}\right)}{1 - \Phi\left(\frac{\lambda \cdot \widehat{s}_i^{inc}(w_i)}{\sqrt{\sigma_{ms}^2 + \sigma_{ws}^2}}\right)} = \frac{\sigma_{ms}^2}{\sqrt{\sigma_{ms}^2 + \sigma_{ws}^2}} \cdot \frac{\phi\left(\frac{(1-\theta)\log(w_i) - \lambda \cdot \bar{V}_{i,f}^{inc}}{\sqrt{\sigma_{ms}^2 + \sigma_{ws}^2}}\right)}{1 - \Phi\left(\frac{(1-\theta)\log(w_i) - \lambda \cdot \bar{V}_{i,f}^{inc}}{\sqrt{\sigma_{ms}^2 + \sigma_{ws}^2}}\right)}. \quad (\text{E.10})$$

## F Assumption TA1 under Gaussian signals

Let us show that TA1 always holds if the signals  $s_{i,f}^{I,ms} + s_{i,f}^{P,ms}$ ,  $f = 1, \dots, F$ , and  $s_i^{I,ws} + s_i^{P,ws}$  are normally distributed. From A1, note first that these signals are independently distributed and have zero mean. Following the notation used for our parametric specification (which relies on many additional parametric restrictions), let  $\sigma_{f,ms}^2$  (resp.  $\sigma_{ws}^2$ ) denote the variance of  $s_{i,f}^{I,ms} + s_{i,f}^{P,ms}$  (resp.  $s_i^{I,ws} + s_i^{P,ws}$ ).

Note that for the incumbent,  $w_{i,f}^{*inc} = u$  is equivalent to  $\lambda \cdot [s_{i,f}^{I,ms} + s_{i,f}^{P,ms} + s_i^{I,ws} + s_i^{P,ws}] = \log(\psi(u)) - \lambda \bar{V}_{i,f}^{inc}$ .

Then according to the same calculation as the one developed to obtain (E.8), we have

$$\log(\mathbb{E}[e^{\lambda \cdot [s_{i,f}^{I,ms} + s_{i,f}^{P,ms}]} | w_{i,f}^{*inc} = u]) = \frac{\sigma_{ws}^2}{\sigma_{f_i^{inc},ms}^2 + \sigma_{ws}^2} \cdot [\log(\psi(u)) - \lambda \cdot \bar{V}_{i,f}^{inc}] + \frac{\sigma_{f_i^{inc},ms}^2}{\sigma_{f_i^{inc},ms}^2 + \sigma_{ws}^2} \cdot \frac{\sigma_{ws}^2}{2}. \quad (\text{F.1})$$

Then according to the notation used for TA1, we get then that

$$\kappa_i(u) = \frac{\sigma_{f_i^{inc},ms}^2}{\sigma_{f_i^{inc},ms}^2 + \sigma_{ws}^2} \cdot \log(\psi(u)) + \frac{\sigma_{ws}^2}{\sigma_{f_i^{inc},ms}^2 + \sigma_{ws}^2} \cdot \bar{V}_{i,f}^{inc} - \frac{\sigma_{f_i^{inc},ms}^2}{\sigma_{f_i^{inc},ms}^2 + \sigma_{ws}^2} \cdot \frac{\sigma_{ws}^2}{2}. \quad (\text{F.2})$$

Given A2, the function  $\log(\psi(\cdot)) : \mathbb{R}_+ \rightarrow \mathbb{R}$  is an increasing bijection. Finally, from (F.2), we obtain that  $\kappa_i(u) : \mathbb{R}_+ \rightarrow \mathbb{R}$  is an increasing bijection.

## G An impossibility result under private values

**Proposition G.1.** *Suppose the auction data is generated by our model (with A1-A2 and TA2) with the given (bijection)  $\psi$  on  $\mathbb{R}_+$  and suppose that the auction never involves an incumbent. For any given bijection  $\psi' \neq \psi$  on  $\mathbb{R}_+$ , we can specify the remaining primitives (i.e. some zero-mean distributions  $G'^{ws}(\cdot | COV)$  and  $\{G_f^{ms}(\cdot | COV)\}_{f=1,\dots,F}$  for any covariate  $COV \in \mathcal{S}_{COV}$ , and the scalars  $\{\lambda \cdot \bar{V}'_{i,f}\}_{f=1,\dots,F}$  for any  $i$  such that  $COV_i \in \mathcal{S}_{COV}$ ) such that our model with these primitives generate the same auction data.*

Proposition G.1 implies that we cannot implement our two-stage strategy. Note that we do not claim here that the CDFs  $\{G_f^{ms}(\cdot | COV)\}_{f=1,\dots,F}$  and the function  $\psi$  can not be identified non-

<sup>62</sup>The variance of  $\bar{s}_{i,f}^{inc}$  is  $\sigma_{ms}^2 + \sigma_{ws}^2$  because the match-specific and worker-specific signals are drawn independently.

parametrically. In particular, we do not rule out that these primitives can be identified by using jointly the bidding and the performance data. The proposition also implies that regarding the auction data, our model is observationally equivalent to a pure private auction model without incentive effect (i.e., with  $s_i^{I,ws} = s_i^{P,ws} = 0$  and  $\tau(\cdot) \equiv 0$  in our setup). This result is reminiscent of [Guerre et al. \(2009\)](#)'s impossibility result according to which the signal distribution and the von Neumann-Morgenstern utility function are not identified.<sup>63</sup>

**Proof of Proposition G.1:** Suppose that the following set of primitives generates the auction data when there is no incumbent: the bijection  $\psi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ , the (zero-mean) CDFs  $G_f^{ms}(\cdot|COV)$  and  $G_f^{ws}(\cdot|COV)$  and the scalars  $\{\lambda \cdot \bar{V}_{i,f}\}_{f=1,\dots,F}$  for any  $i$  with  $COV_i \in \mathcal{S}_{COV}$ .<sup>64</sup>

Let  $A_{i,f} := \lambda \cdot \bar{V}_{i,f} + \log[\mathbb{E}[e^{\lambda \cdot (s_i^{I,ws} + s_i^{P,ws})} | COV_i = COV]]$ . For any  $COV \in \mathcal{S}_{COV}$ , let  $\bar{G}_f(\cdot|COV)$  denote the distribution for  $A_{i,f} + \lambda \cdot (s_{i,f}^{I,ms} + s_{i,f}^{P,ms})$ .

As in the proof of [Proposition 4.1](#), we have then  $H_f^{NI}(\cdot|COV) = \bar{G}_f(\log(\psi(\cdot))|COV)$  where the CDF  $H_f^{NI}(\cdot|COV)$  corresponds to the distribution of the cut-off wage of firm  $f$  when the set of covariates is  $COV$ .

Consider now a bijection  $\psi' : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  and let  $\bar{\psi}' := \log(\psi'(\cdot))$ . Since  $\psi'$  is a bijection, the function  $\bar{\psi}' : \mathbb{R}_+ \rightarrow \mathbb{R}$  is also a bijection. Let  $\bar{G}'_f(\cdot|COV) := \bar{G}_f(\log(\psi(\bar{\psi}'^{-1}(\cdot)))|COV)$ . Consider then the following set of primitives: the function  $\psi'$ , the distribution  $G_f'^{ms}(\cdot|COV)$  defined as the CDF  $\bar{G}'_f(\cdot|COV)$  rescaled to have zero-mean and last the constant  $A'_{i,f}$  equal to the mean of the distribution  $\bar{G}'_f(\cdot|COV)$ . Conditional on  $COV$  and the set of potential participants  $S$  and for an auction without incumbent, we have then that the latter primitives lead to the same distribution  $H_f^{NI}(\cdot|COV)$ ,  $f = 1, \dots, F$  for the cut-off wages as the ones generated by the primitives  $\psi$ ,  $G_f^{ms}(\cdot|COV)$  and  $A_{i,f}$  and thus generate exactly the same auction data (in particular the same distribution for the auction outcome  $(w_i, f_i^w)$ ). **Q.E.D.**

## H Related literature and alternative approaches to identification

### Related literature on identification

The identification argument we have presented does not rely on exogenous variations of the covariates or the reservation wages (as in [Guerre et al. \(2000\)](#)'s seminal contribution and in many subsequent works on the econometrics of auctions (surveyed by [Athey and Haile \(2007\)](#) and [Perrigne and Vuong \(2021\)](#)). Our identification result relies rather on three main restrictions: first, the independence of the signals across bidders; second, the independence between the match-specific and the worker-specific signals; and third, an exclusion restriction stipulating that these signal distributions are identical in the subsamples with and without incumbents. The first independence assumption is needed to apply [Athey and Haile \(2002\)](#), i.e., to identify the cutoff wage distribution of all bidders from the auction outcome. This independence assumption would be sufficient if we were in a setup without incumbents (up to the limits discussed in [Section 4.4](#)). The last two restrictions are crucial to deal with an interdependent value model. Analogs of our second independence restriction are popular to deal with unobserved heterogeneity thanks to a deconvolution argument: in first-price auctions where at least two bids are observed for each given auction, [Li et al. \(2000\)](#) and [Krasnokutskaya \(2011\)](#) identify and estimate private value models where bids are correlated through a common additive independent shock. Restrictions that are similar to our exclusion restriction are used in

<sup>63</sup>There is an analogy between the role played by the function  $\psi$ , reflecting the incentive effect, and by the utility function capturing risk aversion in [Guerre et al. \(2009\)](#), both capturing a departure from quasi-linear preferences.

<sup>64</sup>We do not specify the distribution of the covariates  $COV_i$  and of the set of potential participants  $S_i$  since it does not play any role below.

the literature to test for common values (Haile et al. (2003)), to identify risk-aversion (Lu and Perrigne (2008), Guerre et al. (2009))), or to account for correlated private values (Aradillas-Lopez et al. (2013)). The identification idea consists in exploiting exogenous variations in the environment that induce different bidding functions, while keeping the same signal distributions. For instance, Haile et al. (2003) and Guerre et al. (2009) consider exogenous variations in the number of bidders. To the best of our knowledge, we are the first to exploit exogenous variations in the incumbency status.

**Identification of the auction model:** Alternative approaches could consist in using exogenous variations of the covariates or the reservation wages. For example, the exclusion restriction A3 could be substituted by an assumption stating that the signal distributions do not depend on some covariates (as it is indeed the case in our specification in Section 4.3). As explained in Athey and Haile (2007), such exogenous variations allow identification in a much simpler manner (including when there is a binding reserve price). Intuitively, samples without an incumbent could be substituted by samples with an incumbent having highly unfavorable covariates ( $\bar{V}_{i,f}^{inc}$  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 reservation wages. Binding reservation wages may prevent to identify (non-parametrically) the distributions  $G_f^{ms}$  over their full support. This issue can be solved by assuming that there is sufficient variations in the covariates that are orthogonal to the way private signals are distributed.<sup>65</sup> We stress, however, that our primary goal is to control for wage endogeneity and selection in the performance equation, 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  $G_f^{ms}$  for the realization of  $\lambda \cdot \bar{s}_{i,f}$  such that  $w_{i,f}[NI](\bar{s}_{i,f}) < W_i^r$  because we need to identify the control terms  $CF_i^{ms}[l]$  only up to a constant. Binding reserve prices could nevertheless be helpful for identification, in particular to identify richer models. 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 buyers and sellers but not by the econometrician. We stress that A3 assumes implicitly that reservation wages are exogenous. Dealing with endogenous reservation wage is an interesting avenue for further research.

**Identification of the augmented performance equation:** We have presented the identification of the augmented performance equation (6) for any given  $COV \in \mathcal{S}_{COV}$ , i.e. by exploiting the variation in the auction outcome  $(w_i, f_i^w)$  and the identity of the incumbent  $f_i^{inc}$ .

An alternative path for identifying the augmented performance equation (6) could be to assume that the signal distributions do not depend on the auction shifters  $z_{i,f}$ ,  $f = 1, \dots, F$ , such that the  $\gamma$ -coefficients in A7 do not depend on these auction shifters and that only the terms  $CF_i^{ms}[l]$  and  $CF_i^{ws}[l]$  ( $l = 1, \dots, L$ ) do depend on the auction shifters in (6) which would be a way to guarantee the full rank condition.<sup>66</sup> the auction shifters play the role of instruments.

Another related path for identification could be to assume that the distribution of firm  $f$ 's match-specific signals  $(s_{i,f}^{L,ms}, s_{i,f}^{P,ms})$  depends on  $COV$  only through  $(x_{i,f}, z_{i,f})$ , such that the coefficients  $\gamma_{l,f}^{ms}$  in A7 do not depend on  $(x_{i,f'}, z_{i,f'})$  for  $f' \neq f$ . Relatedly we could assume that the worker-specific signals do not depend on  $COV$  such that the coefficients  $\gamma_l^{ws}$  in A7 do not

<sup>65</sup>E.g. if the distribution of private signals do not depend on  $COV_i$  and if  $\bar{V}_{i,f}$  can be arbitrary large for all  $f$  for some worker  $i$ , then we would be back to the case where the reservation wage is arbitrarily small and the distributions of the aggregate signals  $\lambda \cdot \bar{s}_{i,f}$  are identified on their full support.

<sup>66</sup>We have seen in Section 4.3 and the related section in the SA how the control terms depend on  $z_{i,f}$  and  $z_{i,f}^{inc}$  under our special parametric specification. It is straightforward that this dependence goes beyond our parametric specification provided that  $\bar{V}_{i,f}$  (for each  $f$ ) varies with  $z_{i,f}$ .

depend on  $COV$ . Then the only terms in (6) that depend on  $(x_{i,f'}, z_{i,f'})$  for  $f' \neq f_i^w$  are the control terms  $CF_i^{ms}[l]$  and  $CF_i^{ws}[l]$  ( $l = 1, \dots, L$ ). More precisely, the control terms do not depend on  $(x_{i,f'}, z_{i,f'})$  for  $f' \neq f_i^w$  when  $f_i^{inc} \notin S_i$ . On the contrary, in the sub-sample with an incumbent, some control terms do depend on  $\bar{V}_{i,f'}$  for  $f' \neq f_i^w$  as shown under our special parametric specification. We use then the variation of the match-specific covariates associated to the winning bidder's competitors as a source of identification. The latter source of identification is not specific to auction models but is pervasive to pricing models under imperfect competition where the pricing/bidding strategy of a firm does not depend solely on its own private characteristics (which drive its preferences) but also on the publicly observable characteristics of its opponents. In our model, for example, if there is an incumbent in the auction for a given worker  $i$ , then the bidding strategy of a non-incumbent  $f$  depends on  $\bar{V}_{i,f_i^{inc}}$  and so does also the control terms  $CF_i^{ms}$  when  $f$  wins the auction. Similar dependence also arises in differentiated products market equilibrium as in [Berry et al. \(1995\)](#) where the price of a product can be instrumented by the characteristics of the competing products in the analysis of the market share equation. Nevertheless, note that in the second price or in English auctions under private values, bidders' equilibrium strategies do not depend on their opponents' preferences/characteristics. This is the reason why the control terms  $CF_i^{ms}[l]$  and  $CF_i^{ws}[l]$  ( $l = 1, \dots, L$ ) do not depend on  $(x_{i,f'}, z_{i,f'})$  for  $f' \neq f_i^w$  if  $f_i^{inc} \notin S_i$ .

## I Identification under alternative formats and/or other information structure

The auction format we consider (the English auction) and our assumption on what is observed by the econometrician about bidding behaviors (only the transaction price and the winner's identity are observed) is known to be the most unfavorable case for identification ([Perrigne and Vuong \(2021\)](#)). Under other auction formats, (like the first price auction that is commonly used in procurement), we can easily adapt our methodology and possibly improve it. On the one hand, the independence of the signals would translate immediately into the independence of the bids (contrary to our setup where the auction dynamic assumed to be unobserved by the econometrician can induce correlation among cutoff wages when  $F \geq 3$  which is the reason why we have imposed TA4). On the other hand, 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.<sup>67</sup>

Under private values, the second price auction and the English button auction format with observable exits are strategically equivalent. Nevertheless it is no longer the case with interdependent values, and thus in particular when there is an incumbent. The second price auction is strategically equivalent to the English auction where exits are unobserved. In such a case, the bidding strategy of a bidder can depend only on its own signal and bids are independently distributed which is helpful for identification and would allow to apply [Athey and Haile \(2002\)](#) without using the samples with only two potential participants.

Last, a natural extension would be to consider that in an auction for a worker that has played before, and more generally in an auction for a contract that is renewed (as in [Carnehl and Weiergraeber \(2022\)](#)), not only the incumbent is informed about common value component

---

<sup>67</sup>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 with such form of unobserved heterogeneity.

but also the other bidders (to a lower extent). The main lock to adapt our methodology is the existence of an equilibrium based on monotone bidding function of the aggregate signals. This problem has been solved by [Heumann \(2019\)](#) for the English auction under a Gaussian information structure (and without any incentive effect), but is still an open question in general. How to extent our methodology to environments where multiple bidders receive commonly valued signals is left for further research.

## J Likelihood estimation

The primitives of the auction model are estimated through maximum likelihood and the likelihood function of the auction data is specified in the following way in general: The distribution functions  $G^{ws}(\cdot)$  and  $G_f^{PV}(\cdot)$ ,  $f = 1, \dots, F$ , the increasing bijection  $\psi$  and the functions that maps  $COV_i$  into the term  $\lambda \cdot \bar{V}_{i,f}$  (for  $f = 1, \dots, F$ ), are assumed to be known up to a parameter belonging to  $\mathbb{R}^d$  for some  $d \geq 1$ . Given Corollary 1, these parameters, referred next to as the auction parameters, altogether completely characterize how each firm  $f$  bid as a function of their signal  $\lambda \cdot \bar{s}_{i,f}$ , and this for any worker  $i$  given the set of observable covariates  $COV_i$  and the identity of the incumbent  $f_i^{inc}$ . The cutoff signal functions  $\lambda \cdot w_{i,f}^{-1}(\cdot)$ ,  $\lambda \cdot [w_{i,f}[NI]]^{-1}(\cdot)$ ,  $\lambda \cdot [w_{i,f}[NE]]^{-1}(\cdot)$ ,  $\lambda \cdot [w_{i,f}[DROP]]^{-1}(\cdot; w)$  ( $w \geq W_i^r$ ),  $\lambda \cdot [w_{i,f}[ACT]]^{-1}(w)$  are thus functions of the auction parameters. For notational simplicity we will not explicitly index these cutoff signal functions and the other entities defined below, by the auction parameters.

**Some remarks on the likelihood function that is detailed below:** 1) In writing down the likelihood, one should carefully pick the appropriate density or distribution functions of the cutoff signals associated to observed bidding behavior, 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 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. The full support assumption TA2 guarantees that the auction outcome  $(w_i, f_i^w)$  has full support on  $\mathbb{R}_+ \times S_i$ . Furthermore, in both our simulation protocol and our empirical application, the 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 \(2004\)](#)).

Next we derive the likelihood function of the auction data for observations where there is no RTM, where the set of potential participants is  $\{1, \dots, F\}$ , where an incumbent is present among the potential bidders,<sup>68</sup> and last assuming that the identity of the second-highest bidder is observed by the econometrician (as is the case in our dataset) and denoted next  $f_i^{sh}$  (if there is a single entrant then we have  $f_i^{sh} = 0$ ).

Let  $\mathbb{P}_{i,f_i^{inc}}(w)$  denote the ex ante probability that the incumbent prefers not to be matched with worker  $i$  at wage  $w$  given the value of the auction parameters, i.e. that its valuations is below  $w$ . We thus have  $\mathbb{P}_{i,f_i^{inc}}(w) = G_{f_i^{inc}}(\lambda w_{i,f_i^{inc}}^{-1} | COV_i)$  and  $\mathfrak{l}_{i,f_i^{inc}}(w) = \frac{d\mathbb{P}_{i,f_i^{inc}}(w)}{dw}$  is the corresponding density evaluated at the corresponding cutoff signal of the incumbent. Using the letters B and C associated with the various cases described in our equilibrium analysis, we similarly define  $\mathbb{P}_{i,f}^B(w)$  (resp.  $\mathbb{P}_{i,f}^C(w, w')$  for  $w' \in [W_i^r, w]$ ) as the probability that firm  $f$  ( $f \neq f_i^{inc}$ ) prefers not to be matched with  $i$  at wage  $w$ , conditional on observing that the incumbent has not

<sup>68</sup>The likelihood for the sub-sample without an incumbent (i.e.,  $f_i^{inc} \notin S_i$ ) corresponds to the likelihood for a standard independent private value model and its form can be found in [Paarsch and Hong \(2006\)](#).

entered the auction (resp. has entered and dropped out at wage  $w'$ ). We have thus  $\mathbb{P}_{i,f}^B(w) = G_f^{ms}(\lambda \cdot [w_{i,f}[NE]]^{-1}(w)|COV_i)$  and  $\mathbb{P}_{i,f}^C(w, w') = G_f^{ms}(\lambda \cdot [w_{i,f}[DROP]]^{-1}(w, w')|COV_i)$ . We also define  $\mathbb{P}_{i,f}^D(w)$  as the probability that firm  $f$  ( $f \neq f_i^{inc}$ ) does not wish to be matched with  $i$  at  $w$ , given that the incumbent is still active in the auction at  $w$  and that this firm believes that the incumbent is going to quit instantly at this wage. This probability can hence be written as  $\mathbb{P}_{i,f}^D(w) = G_f^{ms}(\lambda \cdot [w_{i,f}[ACT]]^{-1}(w)|COV_i)$ . Finally, let  $l_{i,f}^k(w) := \frac{d\mathbb{P}_{i,f}^k(w)}{dw}$  for  $k = B, D$  and  $l_{i,f}^C(w, w') := \frac{d\mathbb{P}_{i,f}^C(w, w')}{dw}$ , the associated densities.

For a given value of the parameter vector, the likelihood associated with the event that worker  $i$  remains unmatched conditional on  $\mathcal{A}_i$ , is denoted  $L_i^{unsold}$ . Given Proposition A.1 and the independence assumption A1, we have:

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

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. The likelihood associated with the event that worker  $i$  is matched with firm  $f_i^w$  at  $w_i$ , and the second highest bidder is  $f_i^{sh}$ , conditional on  $\mathcal{A}_i$ , 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 reservation wage.

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

$$L_i^{sold}(w_i, f_i^w, f_i^{sh}) = \prod_{\substack{f=1 \\ f \neq f_i^w}}^F \mathbb{P}_{i,f}^D(W_i^r) \times (1 - \mathbb{P}_{i,f_i^w}(W_i^r)) \text{ if } f_i^w = f_i^{inc},$$

and

$$L_i^{sold}(w_i, f_i^w, f_i^{sh}) = \mathbb{P}_{i,f_i^{inc}}(W_i^r) \times \prod_{\substack{f=1 \\ f \neq f_i^{inc}, f_i^w}}^F \mathbb{P}_{i,f}^B(W_i^r) \times (1 - \mathbb{P}_{i,f_i^w}(W_i^r)) \text{ if } f_i^w \neq f_i^{inc}.$$

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

$$L_i^{sold}(w_i, f_i^w, f_i^{sh}) = \prod_{\substack{f=1 \\ f \neq f_i^w, f_i^{sh}}}^F \mathbb{P}_{i,f}^D(w_i) \times l_{i,f_i^{sh}}^D(w_i) \times (1 - \mathbb{P}_{i,f_i^{inc}}(w_i)) \text{ if } f_i^w = f_i^{inc},$$

$$L_i^{sold}(w_i, f_i^w, f_i^{sh}) = \prod_{\substack{f=1 \\ f \neq f_i^w, f_i^{sh}}}^F \mathbb{P}_{i,f}^D(w_i) \times l_{i,f_i^{sh}}(w_i) \times (1 - \mathbb{P}_{i,f_i^w}(w_i)) \text{ if } f_i^{sh} = f_i^{inc},$$

and



$$\begin{aligned}
L_i^{sold}(w_i, f_i^w, f_i^{sh}) &= \prod_{\substack{f=1 \\ f \neq f_i^w, f_i^{sh}, f_i^{inc}}}^F \mathbb{P}_{i,f}^B(w_i) \times l_{i,f_i^{sh}}^B(w_i) \times (1 - \mathbb{P}_{i,f_i^w}^B(w_i)) \times P_{i,f_i^{inc}}(r_i) \\
&+ \int_{W_i^r}^{w_i} \prod_{\substack{f=1 \\ f \neq f_i^w, f_i^{sh}, f_i^{inc}}}^F \mathbb{P}_{i,f}^C(w_i, w) \times l_{i,f_i^{sh}}^B(w_i, w) \times (1 - \mathbb{P}_{i,f_i^w}^B(w_i, w)) \times l_{i,f_i^{inc}}(w) dw
\end{aligned}$$

if  $f_i^{inc} \neq f_i^w, f_i^{sh}$ .

**Remark:** In our application, we assume that the drop-out wage of the incumbent  $w_i^{inc}$  (when the latter enters the auction and is not the winner) is always observed and we exploit this information to facilitate the maximisation of the likelihood function. Formally, when the incumbent do no longer have a RTM-card, we take  $w_i^{inc}$  as the last wage at which the incumbent has raised its paddle. We then infer from the auction data that  $\lambda \cdot \bar{s}_{i,f_i^{inc}} = \lambda \cdot w_{f_i^{inc}}^{-1}[W_i^r]$  if the incumbent does not enter the auction and that  $\lambda \cdot \bar{s}_{i,f_i^{inc}} = \lambda \cdot w_{f_i^{inc}}^{-1}[w_i^{inc}]$  if the incumbent drop-out at  $w_i^{inc}$ .

## K Monte Carlo: Further Details

### K.1 Simulation Set-up

The aim of the simulation study is to show how our methodology performs in finite samples (we pick  $N = 300$  auctions, but also report results for  $N = 1,000$  auctions in Table K.1). The parameters of our Monte Carlo protocol have been chosen as follows. As in our empirical application, we set the total number of bidders to  $F = 8$ . Throughout, it is assumed that the set of potential participants always 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 among the 8 bidders). In all simulations we fix  $\lambda = 1$ , while  $\tau$  can take the values  $-0.8, -0.5, 0, 0.5$  and  $0.8$  (these are hence also the values taken by  $\theta$ ). We assume that the signals follow a symmetric Gaussian structure as in Section 4.3. In addition, we assume that  $s_{i,f}^{I,ms} = s_{i,f}^{L,ws} = 0$  for all  $i$  and  $f$ , implying that A7 holds with  $L = 1$  and with  $\gamma_1^{ws}(COV) \equiv \gamma^{ws} = 1$  and  $\gamma_{l,f}^{ms}(COV) \equiv \gamma_f^{ms} = 1$ . For all simulations we set  $\sigma_{ms} = 1 - \theta = 1 - \tau$ , while  $\frac{\sigma_{ws}}{1-\tau}$  takes the values 0, 1 and 2 ( $\frac{\sigma_{ws}}{1-\tau} = 0$  corresponds to the private value case considered in Section 4.4). We also assume that  $\beta_{inc} = \alpha_{inc} = 0$ , i.e., the incumbency-indicator  $INC_{i,f}$  is included neither in the performance equation nor in the payoff function  $\mathcal{V}$ . The vectors of covariates  $x_{i,f}$  and  $z_{i,f}$  are both one-dimensional. In line with our empirical application, the vector of firm-worker characteristics  $x_{i,f}$  is assumed to be the same across all bidders, but the auction shifter  $z_{i,f}$  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. We also set  $\bar{V}_{i,f} = x_{i,f} + z_{i,f}$ , which amounts to choosing  $\beta_f^* = 0$ ,  $\beta_x^* = 1$  and  $\beta_z^* = 1$ . This corresponds to the following specifications for the parameter values appearing in payoff function  $\mathcal{V}$ :  $\alpha_f = -\frac{1}{2}$ ,  $\alpha_x = -\tau$  and  $\alpha_z = 1 - \tau$ . Finally, the logarithm of the reservation wage  $W_i^r$  is set to  $\frac{3}{1-\tau}$ . Note that this choice guarantees that the same signal threshold ( $\lambda \cdot \widehat{s}_{i,f}(W_i^r)$ ) as defined in (8)) that makes indifferent between entering or not the auction does not depend on  $\tau$  when  $\sigma_{ws} = 0$ . Nevertheless the probability of sale (reported in the bottom line of Table 1) does depend on  $\tau$  since we have specified  $\sigma_{ms} = 1 - \tau$ .

## K.2 Warp-Speed Monte Carlo

The Warp-Speed method will be described by considering the parameter  $\tau$ . The methodology is strictly the same for any other parameter. From the  $k$ -th replication sample ( $k = 1, \dots, 1,000$ ), 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  $\tau$

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

where  $q_N(0.025)$  and  $q_N(0.975)$  are the 0.025-quantile and 0.975-quantile of the empirical distribution of  $\hat{\tau}_{N,k}^* - \hat{\tau}_{N,k}$ ,  $k = 1, \dots, 1,000$ , respectively. We can now estimate the lower bound (resp. upper bound) of the 95% confidence interval of  $\tau$  by taking the mean over the lower bounds (resp. upper bounds) of  $CI_{N,k}(\tau)$ ,  $k = 1, \dots, 1,000$ . 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  $CI_{N,k}(\tau)$ ,  $k = 1, \dots, 1,000$ , given that the data are generated under a particular value of  $\tau$ . 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.

*Power and Empirical Coverage:* We calculate the power of the test of the null hypothesis  $H_0 : \tau = 0$  (resp.  $H_0 : \beta_x = 0$ ), against the bilateral alternative, as the fraction of times zero does not belong to the 1,000 CIs for  $\tau$  (resp.  $\beta_x$ ), given that the data are generated under a particular value of  $\tau$  (resp.  $\beta_x$ ). Empirical coverage probabilities are reported as the probability that the parameter's true value lies in the confidence interval.

## L 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 and follows the index proposed by the website <https://bit.ly/2CvCB44>.<sup>69</sup> 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 L.1.

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 multiplied by a speciality specific factor. For batting, the factor measures the player’s relative *strike-rate* in the tournament,<sup>70</sup> as the higher the strike rate, the more effective a player is at scoring quickly. For bowlers, the factor measures the bowler’s relative *economy-rate* in the tournament,<sup>71</sup> as the higher the economy rate, the more effective a player 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 either as batsman or as bowlers so that player’s are judged on a per-game basis. We have obtained then two performance measures: one for batting which the performance measure we retain for batsmen and wicket-keepers, and one for bowling which the performance measure we retain for bowlers. For all-rounders (who both bat and bowl), our performance measure is the sum of the two bowling and batting performance measure.

## M Complementary elements on the RTM option

In auctions with RTM-eligibility, the incumbent team may not be able to use the RTM-option because it has exhausted all of its RTM cards. Depending on the number of retained players, each team received from the organizers between 1 to 3 RTM cards. Then a RTM card allowed a team to exert the RTM option for any player from its previous year’s squad. In our data set, the average number of retained player per team is 3 and the average number of RTM cards received by teams at the beginning of the auction sequence is 1.62. Summary statistics related to the RTM option are gathered in Table M.1. We also observe that all teams have exhausted their RTM cards at the end of the auction sequence. In 75 of the 156 auctions with an incumbent, 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.

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

---

<sup>69</sup>Note that the performance measure distinguishes only batting and bowling scores; 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.

<sup>70</sup>A batsman’s strike-rate is defined as the average score of a player per 100 balls faced. Formally, this is equal to  $[100 * (\text{Batsman score} / \# \text{ Balls faced})]$ . The batting factor of a player (either a batsman, a wicket-keeper or an all-rounder) is then equal to his strike-rate divided by the average strike-rate of other batsmen in the tournament.

<sup>71</sup>A bowler’s economy-rate is defined as the average score conceded by a bowler per 6 balls. Formally, this is equal to  $[\text{Bowler Score} / (\# \text{ Balls delivered} / 6)]$ . The bowling factor of a player (either a bowler or an all-rounder) is his economy-rate divided by the average economy-rate of other bowlers in the tournament.

Table K.1: Monte Carlo Results for  $N = 1,000$

| Table 1A: Mean of $\hat{t}$ |       |       |       |       |       |       |       |      |       |      |      |      |      |      |      |
|-----------------------------|-------|-------|-------|-------|-------|-------|-------|------|-------|------|------|------|------|------|------|
| $\tau =$                    | -0.8  |       |       | -0.5  |       |       | 0     |      |       | 0.5  |      |      | 0.8  |      |      |
|                             | 0     | 1     | 2     | 0     | 1     | 2     | 0     | 1    | 2     | 0    | 1    | 2    | 0    | 1    | 2    |
| OLS                         | -0.23 | 0.17  | -0.12 | -0.03 | 0.33  | 0.2   | 0.31  | 0.59 | 0.66  | 0.64 | 0.82 | 0.96 | 0.85 | 0.95 | 1.02 |
| PV                          | -0.78 | -0.47 | -0.4  | -0.49 | -0.21 | -0.09 | 0.01  | 0.19 | 0.36  | 0.5  | 0.61 | 0.72 | 0.79 | 0.85 | 0.89 |
| CV                          | -0.82 | -0.8  | -0.93 | -0.52 | -0.51 | -0.61 | -0.02 | 0    | -0.09 | 0.48 | 0.48 | 0.44 | 0.79 | 0.78 | 0.74 |

| Table 1B: Mean of $\hat{\beta}_x$ |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
|-----------------------------------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| $\tau =$                          | -0.8 |      |      | -0.5 |      |      | 0    |      |      | 0.5  |      |      | 0.8  |      |      |
|                                   | 0    | 1    | 2    | 0    | 1    | 2    | 0    | 1    | 2    | 0    | 1    | 2    | 0    | 1    | 2    |
| OLS                               | 0.43 | 0.09 | 0.39 | 0.52 | 0.22 | 0.38 | 0.68 | 0.46 | 0.43 | 0.84 | 0.73 | 0.63 | 0.93 | 0.89 | 0.83 |
| PV                                | 0.96 | 0.7  | 0.61 | 0.97 | 0.74 | 0.61 | 0.98 | 0.84 | 0.67 | 1    | 0.93 | 0.83 | 1    | 0.98 | 0.93 |
| CV                                | 1.04 | 1.07 | 1.19 | 1.04 | 1.06 | 1.15 | 1.03 | 1.03 | 1.11 | 1.01 | 1.02 | 1.05 | 1.01 | 1.01 | 1.02 |

| Table 1C: Power and CI Bounds $\tau$ |       |       |       |       |       |       |       |       |       |      |      |      |      |      |      |
|--------------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|------|------|------|------|------|------|
| $\tau =$                             | -0.8  |       |       | -0.5  |       |       | 0     |       |       | 0.5  |      |      | 0.8  |      |      |
|                                      | 0     | 1     | 2     | 0     | 1     | 2     | 0     | 1     | 2     | 0    | 1    | 2    | 0    | 1    | 2    |
| Power                                | 1.00  | 1.00  | 1.00  | 1.00  | 0.99  | 1.00  | 0.06  | 0.04  | 0.19  | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Lower Bound                          | -0.99 | -1.04 | -1.14 | -0.67 | -0.72 | -0.82 | -0.17 | -0.20 | -0.28 | 0.32 | 0.30 | 0.26 | 0.54 | 0.50 | 0.52 |
| Upper Bound                          | -0.65 | -0.57 | -0.74 | -0.37 | -0.27 | -0.43 | 0.11  | 0.19  | 0.08  | 0.64 | 0.69 | 0.62 | 1.04 | 1.08 | 0.94 |
| Empirical Coverage                   | 0.96  | 0.95  | 0.72  | 0.93  | 0.96  | 0.76  | 0.94  | 0.96  | 0.81  | 0.95 | 0.95 | 0.92 | 0.96 | 0.93 | 0.87 |

| Table 1D: Power and CI Bounds $\beta_x$ |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
|---|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| $\tau =$                                | -0.8 |      |      | -0.5 |      |      | 0    |      |      | 0.5  |      |      | 0.8  |      |      |
|   | 0    | 1    | 2    | 0    | 1    | 2    | 0    | 1    | 2    | 0    | 1    | 2    | 0    | 1    | 2    |
| 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 | 0.91 | 0.89 | 0.81 | 0.90 | 0.89 | 0.84 | 0.90 | 0.88 | 0.85 | 0.86 | 0.83 | 0.82 | 0.83 |
| Upper Bound                             | 1.22 | 1.34 | 1.47 | 1.20 | 1.31 | 1.43 | 1.17 | 1.22 | 1.33 | 1.15 | 1.18 | 1.23 | 1.18 | 1.20 | 1.21 |
| Empirical Coverage                      | 0.93 | 0.91 | 0.73 | 0.91 | 0.95 | 0.77 | 0.93 | 0.93 | 0.82 | 0.94 | 0.96 | 0.95 | 0.94 | 0.94 | 0.96 |

Table L.1: 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 |

card is costly, then the optimal bidding strategy of the incumbent  $f_i^{inc}$  consists in remaining silent in the auction for worker  $i$  and then in using the RTM card if and only if  $\lambda \cdot \bar{s}_{i,f_i^{inc}}$  is larger than  $\lambda [w_{i,f_i^{inc}}]^{-1}(w)$  where  $w$  is the termination wage in the auction. At a given wage  $w$ , non-incumbents should decide whether or not to remain active as if they knew that the incumbent's aggregate signal  $\bar{s}_{i,f_i^{inc}}$  is below  $[w_{i,f_i^{inc}}]^{-1}(w)$  (which reflects that the incumbent has not used its RTM card at  $w$ ). Then, the cutoff signal of a non-incumbent  $f$  that is indifferent between winning or losing the auction at the wage  $w$ , denoted by  $\hat{s}_{i,f}^{RTM}(w)$ , can be expressed by<sup>72</sup>

$$\lambda \cdot \hat{s}_{i,f}^{RTM}(w) = (1 - \theta) \log(w) - \lambda \bar{V}_{i,f} - \log(\mathbb{E}[e^{\lambda \cdot [s_i^{f,ws} + s_i^{p,ws}]} | \lambda \cdot \bar{s}_{i,f_i^{inc}} < \lambda \cdot w_{i,f_i^{inc}}^{-1}(w)]). \quad (\text{M.1})$$

Under our parametric specification, we obtain then for  $f \neq f_i^{inc}$  (with the same calculation as for case B):

$$\lambda \cdot \hat{s}_{i,f}^{RTM}(w) = (1 - \theta) \log(w) - \lambda \cdot \bar{V}_{i,f} - \log\left(\frac{\Phi\left(\frac{[-\sigma_{ws}^2 + (1-\theta)\log(w) - \lambda \cdot \bar{V}_{i,f_i^{inc}}]}{\sqrt{\sigma_{ms}^2 + \sigma_{ws}^2}}\right)}{\Phi\left(\frac{(1-\theta)\log(w) - \lambda \cdot \bar{V}_{i,f_i^{inc}}}{\sqrt{\sigma_{ms}^2 + \sigma_{ws}^2}}\right)}\right) + \frac{\sigma_{ws}^2}{2} \quad (\text{M.2})$$

Equation (M.2) is the analog of (E.5) where the reservation wage  $W_i^f$  has been replaced by  $w$ .

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, note that the RTM option would simplify a lot the argument: in this subsample, conditional on the observable covariates, the signal of the incumbent is distributed independently of the auction termination wage (which is determined from the competition between non-incumbents which is determined only by their own private signals given the incumbent remain silent until the termination wage is reached). Finally, given our independence assumption A1 ii), the drop-out wages 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.

To sum up, when the incumbent still has a RTM card, then we assume that the incumbent bid as if

<sup>72</sup>Under RTM, the characterization of a separating equilibrium as in Proposition A.1 requires to strengthen a bit TA1 to guarantee that the cutoff signal is increasing in  $w$ . Intuitively, the strength of the informational externality should not be too large.

there were no RTM-card scarcity. Still in order to take into account this scarcity in a reduced form way, then we include as an auction shifter the fact that the player is bought through an RTM through the variable *Incumbent present & no RTM card*.

Table M.1: Summary statistics related to the RTM option

| Variable  | Description  | # Auctions   | Percent of total |
|---|--|--------------|------------------|
| <b># Auctions with incumbent &amp; RTM</b>        | Number of auctions where player's previous team is eligible to use a RTM card          | 75           | 23.66            |
| <i>of which: # Auctions where RTM used</i>        | <i>Number of auctions where player's previous team uses RTM option</i>                 | 13           | 4.10             |
| <b># Auctions with incumbent &amp; no RTM</b>     | Number of auctions where player's previous team is not eligible to use a RTM card      | 81           | 25.55            |
| <i>Bidder purchases and constraints:</i>          |  |              |                  |
|   |  | Mean         | Std. Dev.        |
|   |  | <u>N=317</u> |                  |
| <b># Retained Players</b>                         | Number of players retained by teams before the auctions                                | 3            | 1.85             |
| <b># RTM cards</b>                                | Number of RTM cards received by teams  | 1.62         | 0.74             |
| <b># RTM Bought Players</b>                       | Number of players bought in the auctions through RTM                                   | 1.62         | 0.74             |
| <b># of players where incumbent</b>               | Number of players eligible to be bought back using RTM (per team) if he has a RTM card | 22.37        | 4.56             |
| <b>Auction with incumbent</b>                     | Dummy indicating whether one team is an incumbent (i.e. $\sum_{f=1}^8 I N C_{i,f}$ )   | 0.49         | 0.50             |
| <b>of which RTM</b>                               | Dummy indicating whether the incumbent is eligible to use a RTM card                   | 0.23         | 0.42             |
| <i>RTM eligibility in the performance sample:</i> |  |              |                  |
| <b>Auction with incumbent</b>                     | Dummy indicating whether one team is an incumbent (i.e. $\sum_{f=1}^8 I N C_{i,f}$ )   | 0.74         | 0.43             |
| <b>of which RTM</b>                               | Dummy indicating whether the incumbent is eligible to use a RTM card                   | 0.32         | 0.47             |

Table M.2: Empirical results: specifications with past performance

|   | OLS                     | First Stage                | Second Stage              |
|---|-------------------------|----------------------------|---------------------------|
| Log Wage ( $\tau$ )                                   | 5.89***<br>[2.54, 9.24] |                            | 4.14*<br>[-0.45, 7.79]    |
| <b>Team-player characteristics (<math>x</math>):</b>  |                         |                            |                           |
| Indian  | -3.51<br>[-14.96, 7.94] | -2.57***<br>[-3.41, -1.73] | -1.43<br>[-9.03, 6.06]    |
| Newcomer  | -1.61<br>[-10.19, 6.96] | 0.51*<br>[-0.08, 1.12]     | -3.38<br>[-12.31, 5.71]   |
| Speciality: Batsman                                   | 1.52<br>[-6.87, 9.93]   | 2.24***<br>[1.21, 3.28]    | 3.27<br>[-3.89, 12.44]    |
| Speciality: Bowler                                    | 8.57**<br>[1.57, 15.57] | 2.84***<br>[1.85, 3.84]    | 10.16***<br>[3.15, 18.52] |
| Incumbent present                                     | 1.79<br>[-4.84, 8.43]   | 2.17***<br>[0.97, 3.38]    | 5.39<br>[-3.03, 15.23]    |
| Bidder is incumbent                                   |                         | 2.04***<br>[1.03, 3.06]    |                           |
| Winner is incumbent                                   | -2.37<br>[-8.47, 3.73]  |                            | -16.08*<br>[-31.85, 0.23] |
| Past Available  | -3.51<br>[-14.96, 7.94] | 0.30<br>[-0.54, 1.13]      | -6.21<br>[-17.91, 6.98]   |
| Past Available * Past Performance                     | 0.07<br>[-0.22, 0.37]   | 0.01<br>[-0.02, 0.03]      | 0.19<br>[-0.14, 0.52]     |
| <b>Other Structural Parameters:</b>                   |                         |                            |                           |
| $\gamma^{ms}$   |                         |                            | -0.32<br>[-3.73, 2.67]    |
| $\gamma^{ws}$   |                         |                            | 4.86**<br>[0.68, 35.94]   |
| $\theta$  |                         | 0.76***<br>[0.59, 0.93]    |                           |
| $\sigma_{ms}^* \equiv \frac{\sigma_{ms}}{(1-\theta)}$ |                         | 1.63***<br>[0.77, 3.46]    |                           |
| $\sigma_{ws}^* \equiv \frac{\sigma_{ws}}{(1-\theta)}$ |                         | 2.11***<br>[1.62, 2.74]    |                           |

Note: All specifications account for fixed effects with respect to the set in which the player was auctioned. Column 1 reports OLS estimates of various 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  $\beta_x^*$ ,  $\beta_{inc}^*$ ,  $\theta$ ,  $\sigma_{ms}$ , and  $\sigma_{ws}$ , and 95% CIs based on the asymptotic ML standard errors; the (unreported) auction shifters ( $\beta_z^*$ ) have effects that are similar to those reported in Table 3. Column 3 reports OLS estimates of the parameters in the augmented performance equation (11), and 95% CIs based on a percentile bootstrapped procedure (with 1,000 bootstrapped samples). *Past Available* is a dummy variable taking the value one if past IPL performance is available for the player and *Past Performance* the previous three years' average for a player's past IPL performance. \* indicates significance at 10%; \*\* at 5%; \*\*\* at 1%. The level of significance is based on two-sided tests except for the parameters  $\gamma^{ms}$  and  $\gamma^{ws}$  where it is one-sided given the monotonicity restriction in A7 and 90% CIs are reported in parentheses.