Identification and Estimation of Multi-Period Simple Contracts*

Yonghong An† Shengjie Hong‡ Daiqiang Zhang§

Abstract

This paper provides an econometric framework to analyze two-period simple contracts models where the principal offers the agent to choose between a fixed-price contract and a cost-reimbursement one in each period. We establish nonparametric identification for all model primitives, conditional on that the agent exerts effort. These primitives include agent's cost and disutility functions, distribution of innate costs (type), and parameters that characterize agent's bargaining power and the intertemporal preference. We then propose a consistent estimation procedure. In our empirical study, using the data on transport procurement contracts in France, we find strong evidence that the agent's optimal effort is monotonic in its innate cost, rather than being a constant as would be implied by linear cost functions. A counterfactual analysis evidences the importance of this finding by showing that welfare implications under monotone optimal effort differ a lot from those under constant optimal effort.

Keywords: Simple contracts, multi-period, measurement error, nonparametric identification.
JEL: C14, D82.

*Hong gratefully acknowledges support from the National Natural Science Foundation of China (Grant No. 71703078).

†Department of Economics, Texas A&M University, College Station, TX 77843; email: y.an@tamu.edu.
‡SEM, Tsinghua University, Beijing, China 100084; email: hongshj@sem.tsinghua.edu.cn.
§Department of Economics, University at Albany, State University of New York, Albany, NY 12222; email: dzhang6@albany.edu.

1 Introduction

Due to the fundamental role it plays in the studies of informational asymmetries and incentives, contract theory has attracted much attention from economists during the past three decades. One branch of the contracts are the *complex optimal contracts* in the spirit of Laffont and Tirole (1986) where the optimal payment to agent is a nonlinear function of both agent's unobserved type and their observed cost. Recent studies suggest that another branch, the *simple menus of contracts* which oftentimes specify the payment only as a function of the agent's observed cost or even as a constant, could be more useful in practice (e.g., Bajari and Tadelis, 2001). Theoretical and empirical evidence show that these simple menus could capture a substantial proportion of the surplus that complex optimal nonlinear contracts would achieve (Rogerson, 2003; D'Haultfoeuille and Février, 2015).

Despite the theoretic importance of simple contracts and their wide usage in various sectors in practice, econometric analyses on this large class of contracts are largely missing in the literature. In this paper we provide the first set of positive results on identification of multi-period simple contracts by focusing on the "fixed-price-cost-reimbursement (FPCR)" menu. This menu consists of a fixed-price (FP) contract, in which the payment to the agent is a fixed price, regardless of the agent's realized cost; and a cost-reimbursement (CR) contract, in which the agent is reimbursed exactly for the realized cost. The FPCR menu is widely used in practice. For example, FP contracts or CR contracts are commonly employed by the U.S. Department of Defense (Rogerson, 1992). Many local authorities in France also use the FPCR menu to contract with operators to provide the transport service. Other examples include the Indian customized software industry (Banerjee and Duflo, 2000), the U.S. Air Force engine procurement (Bajari and Tadelis, 2001), the offshore software industry (Gopal et al., 2003), and among others.

A multi-period FPCR contract may be implemented in two forms: a contract under renegotiation, which allows the principal and the agent to renegotiate on the initial contract at certain time during the implementation, or a contract under commitment, which prohibits any adjustment of initial contract during implementation. Since the equilibrium outcome of a multi-period contract under commitment is identical to that of a one-period (static) contract, we focus on the more flexible contracts under renegotiation.

We first construct a theoretical model for two-period (dynamic) FPCR contracts under renegotiation by extending the work of Gagnepain et al. (2013), which provides a simple framework for dynamic FPCR contracts. Briefly, our theoretical model is set up as follows: At the beginning of the first period, an agent chooses the most profitable contract from the

FPCR menu provided by the principal for two periods. The agent is allowed to renegotiate with the principal and make changes to his initial choice at the end of the first period if the contract is renegotiable. At the equilibrium, the range of agent's types can be distinguished into three segments, with agents whose types belong to a same segment making the same choice of contracts: Most efficient agents choose FP contracts in both period; Medium efficient agents choose a CR contract in the first period, followed by a FP contract in the second period; And least efficient agents choose CR contracts in both periods.

We then provide constructive strategies to identify the structural elements, which take multiple steps as follows: (i) By adopting the recently developed methodology in measurement errors (Schennach and Hu, 2013), we recover the distribution of the unobserved optimal effort exerted by agent from the joint distribution of two observable covariates correlated with the effort. Further, the one-to-one mapping between agent's observed cost and optimal effort implied by the equilibrium conditions enables us to back out the optimal effort corresponding to each of the observed cost; (ii) Relying on an exclusion restriction, i.e., the existence of some exclusion variable that directly affect the optimal effort but is independent of innate costs (types), we identify the cost structure of agent. The identification is achieved by exploiting the heterogeneous quantiles of the realized cost conditional on the exclusion variables while the quantiles of the innate cost remain the same; (iii) We recover the innate costs from the identified cost function as well as the structural link between innate cost and optimal effort. Consequently the distribution of the innate cost on the support associated with fixed-price contracts can be recovered. We then employ the structural elements identified above and the observed payment to the agent to recover parameters that characterize agent's bargaining power and intertemporal preference.

Due to the fact that the optimal effort associated with CR contracts is always zero, it is impossible to identify the cost function on the part of its domain that corresponds to CR contracts without imposing additional assumptions. We show that once the cost function is parametrized, we are able to nonparametrically identify all other model primitives on their full support given the identification of cost parameters. Based on the argument of identification, we propose a feasible semiparametric procedure to estimate the model components.

Our identification strategy is potentially applicable to several variations of FPCR menus. A leading case is a linear cost sharing-cost reimbursement (LCSCR) menu where the fixed-price contract is replaced by a linear cost-sharing (LCS) contract (e.g., Chu and Sappington, 2007). The applicability of our identification result to LCSCR menu follows from the fact that LCS contracts provide additional variation of payments compared with FP contracts. Our strategy also readily carries over to one-period (static) FPCR contracts and two-period contracts with commitment. This is because the equilibrium outcomes in a one-period setting is similar to that in the two-period setting with renegotiation, in the sense that they provide similar information for identification, and the two-period contracts with commitment is just twice-repeated version of equilibrium in a one-period (static) setting, as proved in Laffont and Tirole (1990). Last but not least, our identification results for two-period FPCR contracts can be potentially extended to more flexible menus where the observed costs may be linked with agent's unobserved type and effort.

Finally, we apply our method to study the dynamic transport procurement contracts in France. The objective of the empirical study is to estimate the model primitives and to test the convexity of the cost function, which induces monotone optimal effort, rather than constant optimal effort induced by a linear cost function which is widely adopted in the related literature (e.g., Laffont and Tirole, 1988, 1990, Rogerson, 2003, and Gagnepain et al., 2013, among many others). Our estimates suggest that the cost function of agent is convex in innate cost. Moreover, we utilize these estimates to conduct a counterfactual analysis on the welfare comparison between FPCR contracts under monotone optimal effort and constant optimal effort. The counterfactual analysis shows that in French transport industry the social welfare of FPCR contracts under monotone optimal effort is 14.6 million euros more than that under constant optimal effort. An important insight from these empirical findings is that, when investigating the efficiency and welfare implications of simple contracts (either as researchers or policy makers), it is crucial to take into account the monotonicity of optimal effort determined by the functional form of agent's cost function.

This paper contributes to a large literature on the econometrics of contract models. In this literature, studies on identification are still limited. Notably, Perrigne and Vuong (2011) establish nonparametric identification of a static *complex* contract model tailored from the seminal paper Laffont and Tirole (1986). More recently, D'Haultfoeuille and Février (2015) show *partial* nonparametric identification of simple compensation contracts using exogenous variations of contracts. Our study differs from both. The identification argument in Perrigne and Vuong (2011) does not apply to the simple contracts considered in this paper. This is mainly because the one-to-one mapping between the observed price of the product and the private type of the agent, a key to identification in Perrigne and Vuong (2011), is unavailable for simple contracts where such a payment relationship does not exist. Compared with D'Haultfoeuille and Février (2015), our paper differs in both model setup and identification strategies.[1] In a broader view, our identification result is related to the econometric analysis on a richer class of moral hazard and adverse selection models where an agent with a continuum of types is only offered a menu of a few simple contracts by the principal, for example the insurance models studied by Aryal et al. (2010) and the nonlinear pricing models studied by Luo et al. (2018). In addition, identifying structural models through measurement errors has been widely used in the literature. See, for instance, Hu (2017) for a recent survey. Nevertheless, to the best of our knowledge, our paper is the first to employ the results in measurement errors to identify contract models. Our empirical study contributes to a growing literature on empirical industrial organization (IO) studying contracts. Our study is most closely related to Gagnepain et al. (2013), who use the same source of data to evaluate the negotiation cost of FPCR contracts under the assumption of a linear cost function.

The rest of the paper is organized as follows. Section 2 presents the general dynamic FPCR model. Section 3 establishes the main identification results. And Section 4 provides a feasible estimation procedure. Section 5 analyzes the French transport procurement contracts. Section 6 concludes. Proofs are collected in the appendix.

2 The Model

2.1 Basic setup

A risk-neutral principal wishes to procure a project from a risk-neutral agent by offering a two-period menu consisting of two types of contracts in each period: a fixed-price (FP) contract in which the payment is dependent upon the agent's realized cost; and a cost-reimbursement (CR) contract in which the agent is reimbursed exactly for the realized cost. The menu of contracts consists of one FP contract and one CR contract is called fixed-price-cost-reimbursement (FPCR) menu. An agent's innate cost (or "type") θ is a random draw from a cumulative distribution function $F(\cdot)$, with a density $f(\cdot)$ on a support $[\underline{\theta}, \bar{\theta}] \subset \mathbb{R}$. The agent observes its private value of θ, and the principal only has the knowledge of $F(\cdot)$. The cost structure of realizing the project is

$$c_t = H(\theta - e_t),\ t \in \{1, 2\},$$

[1] In D'Haultfoeuille and Février (2015), agent's disutility function of effort is not explicitly modeled, in the sense of which their setting deviate from standard contract models. In contrast, Our paper focus on a large class of standard regulatory contracts. D'Haultfoeuille and Février (2015) provide an informative partial identification result. We aim at developing regularity conditions for point identification.

where c_t is the realized cost in period t; $H(\cdot)$ is a general cost function; the innate cost θ represents the agent's management and production skills, which is invariant during two periods, and the private effort $e_t \geq 0$ captures agent's actions taken to reduce cost c_t. The exerted effort e_t incurs some disutility according to a disutility function $\psi(e_t)$. Both type θ and effort e_t are private information to the agent.

Given the principal's payment schedule $q_t, t = 1, 2$, the agent maximizes the following intertemporal profit by choosing a two-period contract from the FPCR menu and then exerting the optimal effort e_t,

$$u = ru_1 + (1 - r)u_2, \tag{1}$$

where we suppress the arguments of u, u_1 and u_2 for ease of exposition. The weight r is defined as $r = 1/(1 + \delta)$ with δ being the discount factor, it is a measure of the relative length or the relative importance of the first period.[2] u_t is the the information rent (profit) of the agent with type θ in period $t = 1, 2$:

$$u_t \equiv q_t - c_t - \psi(e_t) = q_t - H(\theta - e_t) - \psi(e_t). \tag{2}$$

The principal designs the optimal menu of contracts by specifying the two-period fixed prices (q_1, q_2) for FP contracts to maximize the expected social welfare

$$\pi = \int \tilde{\pi}(\theta) dF(\theta),$$

where $\tilde{\pi}(\theta)$ is the social welfare generated by the agent with innate cost θ:

$$\tilde{\pi}(\theta) \equiv S - (1 + \lambda)[rq_1 + (1 - r)q_2] + \alpha[ru_1 + (1 - r)u_2], \tag{3}$$

In the definition above, the dependence of the right side of (3) on θ is through the dependence of both u_1 and u_2 on θ. S is the gross surplus generated by the procured service and assumed to be sufficiently large to guarantee the desirability of the project. The cost of public funds $\lambda > 0$ captures some dead-weight loss due to a distortionary taxation for raising subsidies with the principal's intertemporal payment $rq_1 + (1 - r)q_2$. The parameter α is the weight assigned to firms' profits by the principal (Baron and Myerson, 1982; Baron, 1988), which can reflect the extent of the political pressure imposed by the agent on the political principals (Faure-Grimaud and Martimort, 2003) and therefore

[2] The relationship between the payoffs in the first and second period is determined by the intertemporal weight r. In this sense δ not only reflects the discount factor, but, more in general, the relationship between the payoffs in the two periods. This interpretation follows from Laffont and Tirole (1990), who assume $\delta \in (0, \infty)$.

can be interpreted as firm's bargaining power against the principal in negotiation. We maintain in our paper that $\alpha < 1 + \lambda$, which captures the relevant trade-off between extracting profit and inducing efficient cost-reducing effort. Intuitively, the optimal menu of contracts offered by the principal trades off efficiency and rent extraction: FP contracts with large subsidy would induce the first-best effort while leaving much information rent to more efficient agent; CR contracts, however, nullify this rent without any incentive for agent to make cost-reducing effort.

2.2 The problems of principal and agent

Based on the setup above, we analyze the principal and agent's problem. We maintain that both $H(\cdot)$ and $\psi(\cdot)$ are twice continuously differentiable and that integration and differentiation can be interchanged. For a generic function $a(\cdot)$ with more than one argument, we denote its derivative with respect to the k-th argument by $a_k(\cdot)$. We impose the following standard assumption in contract literature (e.g., Laffont and Tirole, 1993).

Assumption 1 *(i) Agent's cost function $H(\cdot)$ satisfies $H(\cdot) \geq 0$, $H'(\cdot) > 0$, $H''(\cdot) > 0$. (ii) Agent's disutility function satisfies $\psi(\cdot) \geq 0$, $\psi'(\cdot) > 0$, $\psi''(\cdot) > 0$, $\psi(0) = 0$.*

Recall that in each period, an agent with type θ makes a choice between FP and CR contracts, and then exerts effort to fulfill the contract by maximizing profit specified in (2). Let e^* denote the optimal effort of the agent. If the agent chooses a CR contract, he is only reimbursed the realized cost. We can show under Assumption 1 that in a CR contract 1, $e^* = 0$ and the cost is

$$c = H(\theta). \tag{4}$$

If the agent chooses a FP contract, he will exert a type-dependent optimal effort $e^*(\theta) > 0$, which satisfies the first order condition

$$H'(\theta - e^*) = \psi'(e^*). \tag{5}$$

The realized cost of the agent is

$$c = H(\theta - e^*(\theta)). \tag{6}$$

We summarize the results in the following lemma.

Lemma 1 *Under a CR contract, an agent with type θ exerts no effort and the cost is $c = H(\theta)$. If the agent chooses an FP contract, his optimal effort $e^*(\theta)$ is strictly increasing in type θ and $0 < de^*/d\theta < 1$.*

Denote by $\underline{e}$ the proper lower bound of effort levels associated with all FP contracts. An implication of Lemma 1 is $\underline{e} = e^*(\underline{\theta})$, i.e. $\underline{e}$ is the optimal effort corresponding to $\underline{\theta}$.

Next, we analyze the two-period contract with renegotiation by using the renegotiation-proof principle in the related literature.[3] Under renegotiation the principal provides three options for an agent: A two-period FP contract (FF contract) characterized by the fixed prices in two periods: (b_1, b_2); a first-period CR contract followed by a second-period FP contract $(H(\cdot), b_3)$, denoted by CF contract, where b_3 is the fixed-price in the second period; and a two-period CR contract $(H(\cdot), H(\cdot))$, denoted by CC contract, where $H(\theta)$ indicates the cost in a CR contract.[4] By using the renegotiation-proof principle, we obtain the following proposition that characterizes the equilibrium outcome of the renegotiation-proof menu of contracts (b_1, b_2, b_3) :

Proposition 1 *Under Assumption 1, the optimal renegotiation-proof menu of two-period contracts* (b_1, b_2, b_3) *satisfy* $b_1 = b_2 \equiv \underline{b} < b_3 \equiv \bar{b}$ *with two cut-off types* (θ_l, θ_u) *such that* $\underline{\theta} \leq \theta_l < \theta_u \leq \bar{\theta}$ *and*

$$\bar{b} = H(\theta_u - e^*(\theta_u)) + \psi(e(\theta_u)),$$

$$\underline{b} = r[H(\theta_l - e^*(\theta_l)) + \psi(e^*(\theta_l))] + (1 - r)\bar{b},$$

$$\left(1 - \frac{\alpha}{1 + \lambda}\right) \frac{F(\theta_u) - F(\theta_l)}{f(\theta_u)} = \frac{H(\theta_u) - \bar{b}}{H'(\theta_u - e^*(\theta_u))}. \tag{7}$$

The most efficient types within the lower subinterval $[\underline{\theta}, \theta_l]$ *choose FP contracts in both periods; the intermediate efficient ones within the intermediate subinterval* $(\theta_l, \theta_u]$ *choose CR and FP contracts in the two periods, respectively; and the least efficient ones within the larger subinterval* $(\theta_u, \bar{\theta}]$ *choose CR contracts in both periods.*

The intuition for $\underline{b} < \bar{b}$ can be ascribed to the fact that fixed prices must be raised sufficiently to induce those intermediate efficient firms with the initial choice of CR contracts to switch to FP contracts when the information on the agent's innate cost is revealed after the first period, whereas most efficient firms would choose FP contracts at the beginning even they are paid a relatively low fixed price $\underline{b}$ in both periods.

[3] Renegotiation-proof principal: if a contract is a perfect Bayesian equilibrium in which renegotiation occurs in equilibrium, then there exists a renegotiation-proof contract that achieves the same outcome. In other words, any long-term agreement which is renegotiable could be replaced by another long-term contract with a second-period continuation equal to the renegotiated offer.

[4] Note that the option of CF contract arises from the fact that agent may renegotiate to change the choice of contract made at the beginning of the first period. The choice consisting of a first-period FP fixed price contract followed by a second-period CR contract is never optimal for an agent. This is because the profit from an FP contract is strictly positive almost surely while the profit with CR contracts is always zero.

3 Identification

This section presents identification of the model primitives for the optimal renegotiation-proof menu of two-period contracts. In what follows we denote random variables by upper-case letters and their realized values by lower-case letters. For each contract, the data report realized cost C, two fixed prices $(\underline{B}, \overline{B})$, and two binary variables $D^{(1)}$, $D^{(2)}$, with $D^{(1)} = 1$ and $D^{(2)} = 1$ indicating agent's choice of FF and CF contract, respectively. Consequently, the binary variable $1 - D^{(1)} - D^{(2)}$ indicates the choice of CC contract. We also observe a vector of variables Z that summarize the characteristics of the principal, agent, and contract. Our identification result applies conditional on Z. While presenting our method in this section, we suppress Z in the notation for simplicity. Let $\mathcal{S} \equiv [F(\cdot), H(\cdot), \psi(\cdot), \alpha/(1+\lambda), r]$ denote our model primitives to be identified. Note that the parameters α and λ cannot be separately identified because only the ratio $\alpha/(1+\lambda)$ matters for the optimal menu of contract designed by the principal (r.f. (7)). We maintain that the observed data are generated from the model primitives $\mathcal{S}$ and the equilibrium conditions presented in the preceding section are satisfied .

The model is identified up to a monotone transformation of the innate cost θ. The lemma below shows that an alternative model structure $\widetilde{\mathcal{S}} \equiv [\widetilde{F}, \widetilde{H}, \widetilde{\psi}, \alpha/(1+\lambda), r]$ where $\widetilde{F}(\cdot) = F(\cdot/\xi_1)$, $\widetilde{H}(\cdot) = H(\cdot/\xi_1)$, and $\widetilde{\psi}(\cdot) = \psi(\cdot/\xi_1)$ for some scalar $\xi_1 > 0$, is observationally equivalent to the structure $\mathcal{S}$ in the sense that they both lead to the same joint distribution of $(C, \underline{B}, \overline{B}, D^{(1)}, D^{(2)})$.

Lemma 2 ***(Observational Equivalence)*** *Suppose two structures $\mathcal{S} \equiv [F, H, \psi, \alpha/(1+\lambda), r]$ and $\widetilde{\mathcal{S}} \equiv [\widetilde{F}, \widetilde{H}, \widetilde{\psi}, \alpha/(1+\lambda), r]$ both satisfy Assumption 1, then $\mathcal{S}$ and $\widetilde{\mathcal{S}}$ are observationally equivalent.*

Intuitively, the observational equivalence arises from the fact that both the cost function $H(\cdot)$ and its argument (type) θ are unobservables. A linear transformation $\widetilde{\theta} = \xi_1\theta$ of the type θ together with an appropriate transformation of $H(\cdot)$ leads to the same realized cost. Accordingly, we may adjust other model primitives to rationalize other observables $(\underline{B}, \overline{B}, D^{(1)}, D^{(2)})$.

We will normalize the lower bound $\underline{\theta}$ of the type support, which turns out to be sufficient to rule out observational equivalence scenario as described in Lemma 2 and the upper bound $\overline{\theta}$ will be identified whenever necessary.[5] Consequently, $[\underline{\theta}, \bar{\theta}]$ is divided into three subintervals $[\underline{\theta}, \theta_l]$, $[\theta_l, \theta_u]$ and $[\theta_u, \bar{\theta}]$, where an agent with type belonging to $[\underline{\theta}, \theta_l]$ or

[5] Perrigne and Vuong (2011, 2012) use similar normalizations in the nonparametric identification of a static contract model with adverse selection and moral hazard.

$[\theta_l, \theta_u]$ would choose the FF contract or the CF contract, respectively, and an agent with type belonging to the last subinterval would choose the CC contract. Since FP contracts provide high incentives for cost-reducing effort and CR contracts have no incentive for effort, these two types of contracts induce different equilibrium conditions on the optimal effort. That is, the optimal effort with FP contracts depends on the agent's type while the optimal effort with CR contracts is always zero for any type of agent. Hence, our identification strategy will be constructed in terms of FP contracts and CR contracts, respectively.

3.1 Identification of optimal effort

The first challenge of identification concerns the unobserved optimal effort. Although the static monopoly contract model in Perrigne and Vuong (2011) also involves the unobserved effort, their identifying strategy does not apply to our model because other than the observed cost, more observables (i.e., output and price) in their model are available, which is helpful for identification. By contrast, the directly related data in our model only consist of one observed continuous random cost variable and two binary choice variables. Acknowledging this difficulty, we adopt a newly developed method in measurement error, i.e., Schennach and Hu (2013), to back out the distribution of the optimal effort nonparametrically.

Assumption 2 *There exist two measurements of effort E^*, X and Y, that satisfy:*

$$\begin{aligned} X &= E^* + V_1, \\ Y &= m(E^*) + V_2, \end{aligned} \tag{8}$$

where $m(\cdot)$ is an unknown function and E^*, V_1, and V_2 are mutually independent with $\mathbb{E}V_1 = \mathbb{E}V_2 = 0$.

In Assumption 2, the first measurement X can be understood as a normalization of the level of the unobserved optimal effort and the second measurement Y can be chosen flexibly since $m(\cdot)$ can be a general function. This assumption is much less restrictive than the existence of double measurements of a latent variable needed for the identification of many structural models, e.g., in Li (2002), where $m(\cdot)$ must be an identity function. The optimal effort or "hidden action" in contract theory is generally unverifiable, but agent's effort-related performance is oftentimes measurable which can thus be used as measurements for the optimal effort. For example, Cicala (2015) discusses the plausibility of employing cost-related variables to infer agent's effort. In the empirical analysis of transportation procurement contracts in France, X and Y are chosen to be the share of drivers among all the employees (employees consist of drivers and engineers) and the labor fee, respectively.

Theorem 1 in Schennach and Hu (2013) show nonparametric identification of $m(\cdot)$ and $F_{E^*}(\cdot)$ are both nonparametrically from the joint distribution of X and Y except in some special cases. These cases would impose little restrictions to our model, as discussed in Appendix. In addition to Assumption 2, several additional regularity conditions are required for the identification. Nevertheless, these conditions impose no further meaningful restrictions to our model primitives. The main idea of the identifying strategy is to investigate the higher-order moments (characteristic functions) for the joint distribution of X and Y, which provides sufficient information to secure identification of the function $m(\cdot)$ and distribution $F_{E^*}(\cdot)$.

Recall that the optimal effort is increasing in agent's innate cost $0 < e^{*'}(\theta) < 1$ for any $\theta \in [\underline{\theta}, \theta_u]$. The inverse function theorem implies that the innate cost is a strictly increasing function of the optimal effort with $\theta'(e^*) > 1$. Consequently, the argument $\theta - e^*(\theta)$ in $H(\cdot)$ is strictly increasing in the optimal effort e^* since $d(\theta - e^*)/de^* = \theta'(e^*) - 1 > 0$. Let $L(e^*) \equiv \theta(e^*) - e^*$. Then the observed cost for an agent who choose FF or CF contract, $c = H(\theta(e^*) - e^*)$ can be rewritten as

$$c = H(L(e^*)) \equiv \widetilde{L}(e^*),$$

with $\widetilde{L}'(\cdot) = H'(\cdot)L'(\cdot) > 0$. A direct implication of this result is the existence of a one-to-one mapping between the cost and the optimal effort, and the mapping enables us to obtain the following important structural link:

$$F_C(c) = \Pr(\widetilde{L}(e^*) \leq c) = \Pr(e^* \leq \widetilde{L}^{-1}(c)) = F_{E^*}(\widetilde{L}^{-1}(c)) = F_{E^*}(e^*), \tag{9}$$

where $F_C(\cdot)$ is the cumulative distribution function of cost C. The relationship above in turn implies that

$$e^* = F_{E^*}^{-1}(F_C(c)), \quad c \in [\underline{c}, c_u], \tag{10}$$

where $\underline{c} \equiv H(\underline{\theta} - e^*(\underline{\theta}))$ and $c_u \equiv H(\theta_u - e^*(\theta_u))$. Therefore, observing the realized cost c from any contract, we can recover the corresponding optimal effort e^* as in (10) for FP contracts, or as zero for CR contracts. We summarize this identification result in the following proposition.

Proposition 2 *Under Assumptions 1-2, the optimal effort e^* associated with any given realized cost c is identified.*

3.2 Identification of cost function

The identification of $H(\cdot)$ relies on the recovered optimal effort e^* and the structural link between observed cost c and the cost function $H(\cdot)$ specified in (6).

$$c = H(\theta - e^*(\theta)).$$

In addition, our identification strategy requires an exclusion variable (denoted by W) that is independent from agent's type θ, yet affects the disutility from exerting any given effort level e as $\psi(e, W)$. The variation of W causes quantiles of cost to change while the corresponding quantiles of type are unchanged. This allows us to identify the cost function. As we will show, a binary W would be sufficient to achieve identification. Therefore, we assume $W \in \{\varpi_1, \varpi_2\}$ without loss of generality. Recall that the optimal effort e^* of an agent who chooses FP contract is determined by the first-order-condition $H'(\theta - e^*) = \psi'(e^*)$, thus in general e^* is a function of W. Consequently, the cutoff values θ_l and θ_u also depend on W. For simplicity of exposition, we define $\psi_j(\cdot) \equiv \psi(\cdot, \varpi_j)$, $\psi'_j(e) \equiv \partial\psi(e, \varpi_j)/\partial e$, $C_j(\theta) = H(\theta, e^*(\theta, \varpi_j))$, and $e_j(\theta) = e^*(\theta, \varpi_j)$ for $j = 1, 2$. Proposition 1 states that only the agent with type on $[\underline{\theta}, \theta_u]$ chooses FF or CF contracts (i.e. FP contract at least once in the two periods), and the realized costs is on $[\underline{c}, c_u]$. When conditional on W, we introduce the following notations: an agent with $W = \varpi_j$ would choose FF or CF contracts only if $\theta \in [\underline{\theta}, \theta_u^j]$, with $e^* \in [e_j(\underline{\theta}), e_u^j]$ and the resulting $c \in [c_j(\underline{\theta}), c_u^j]$, where $(\theta_u^j, e_u^j, c_u^j)$ are the counterparts (θ_u, e_u, c_u) with $W = \varpi_j$ for $j = 1, 2$.

Assumption 3 *There exists an observable variable W such that: (i) θ is independent of W. (ii) $\psi(\cdot)$ depends on W such that $\psi_1(e_1(\underline{\theta})) = \psi_2(e_2(\underline{\theta}))$, $\psi'_1(e_1(\underline{\theta})) = \psi'_2(e_2(\underline{\theta}))$, and $\psi'_1(e) > \psi'_2(e)$ for any $e > max\{e_1(\underline{\theta}), e_2(\underline{\theta})\}$.*

Assumption 3 requires the existence of an exclusion variable, such that the distribution of the innate cost does not rely on this variable, but the disutility function does, which captures the heterogeneity of the disutility across agents when cost-reducing effort is exerted by agents who have the same innate productivity (type). The exclusion variable W can be a component of the characteristics vector Z we introduced earlier. In our empirical application of transport procurement contracts, a suitable choice of such a variable can be a dummy variable indicating whether a firm is publicly or privately owned: the ex ante managerial ability of firms does not depend on their ownership, while the disutility of exerting cost-reducing effort may vary with firm's ownership, which can be partly explained by the firm's internal structure (Gagnepain et al., 2013).

Part (ii) of Assumption 3 implies that $\underline{e}_1 = \underline{e}_2 = \underline{e}$. This is due to the first order condition (5) evaluated at $\underline{\theta}$:

$$H'(\underline{\theta} - e_j(\underline{\theta})) = \psi_j'(e_j(\underline{\theta})),$$

as well as the strict monotonicity of $H'(\cdot)$. This property states that the most efficient agent exerts the same (the minimum) level of effort, regardless his characteristics indicated by W. Note that this assumption is empirically testable because $e_1(\underline{\theta}) = e_2(\underline{\theta})$ is equivalent to $c_1(\underline{\theta}) = c_2(\underline{\theta})$, i.e., the lower bound of cost with w_1 equals that with w_2. Consequently, the last condition in part (ii) reduces to $\psi_1'(e) > \psi_2'(e)$ for any $e > \underline{e}$. Part (ii) imposes a single crossing condition on the relationship between W and the marginal disutility of effort. The single or multiple crossing property plays a crucial role in economic theory, and is widely used in the identification literature (e.g. Chesher (2003), Chernozhukov and Hansen (2005), Heckman et al. (2010), Torgovitsky (2015)). Part (ii) of Assumption 3 can be relaxed: given that ψ_1' and ψ_2' across at least once (the intersection point is not necessarily at the lower bound) on the support of effort, our identification result still holds. We discuss the result as an extension later in this section.

Now we provide the key insights for identifying the cost function $H(\cdot)$. For those types associated with effort exerted by the agent, we have

$$c_j(\theta) = H(\theta - e_j(\theta)), \;\; \theta \in [\underline{\theta}, \theta_u^j], \;\; j = 1, 2, \tag{11}$$

where the realized cost C_j is observable and the effort E_j is recovered in Proposition 2. Considering that $H(\cdot)$ is strictly increasing, the equation above implies that

$$\theta = H^{-1}(c_j(\theta)) + e_j(\theta), \;\; \theta \in [\underline{\theta}, \theta_u^j], \;\; j = 1, 2. \tag{12}$$

Denote by $\theta(\tau)$, $E_j(\tau)$, and $C_j(\tau)$ the τ-th quantiles of θ, $e_j(\theta)$, and $c_j(\theta)$, respectively. Our identification strategy is to exploit the condition that the distribution of θ is independent of W to establish the link between $e_1(\theta)$ and $e_2(\theta)$ which correspond to a same θ. Specifically, for a given quantile τ of θ, $\theta(\tau)$, the corresponding quantiles of $e_1(\cdot)$ and $e_2(\cdot)$ differ, and so do the quantiles for $c_1(\cdot)$ and $c_2(\cdot)$ (for more details, see the proof of Proposition 3 in the Appendix.) And this is partly because θ_u^1 can be different from θ_u^2. Without loss of generality, we assume $\theta_u^1 < \theta_u^2 \equiv \theta^*$, which is equivalent to a testable statement that the lower bound of cost for CC contract under $W = \varpi_1$ is smaller than that under $W = \varpi_2$. We further define

$$p_j \equiv \Pr(\underline{\theta} < \theta < \theta_u^j) \in (0, 1),$$

which measures the proportion of firms who choose FP contract at least in one period. It follows from $\theta_u^1 < \theta_u^2$ that $p_1 < p_2$. It is easy to show that p_j is directly identified as,

$$p_j = \Pr(\underline{\theta} < \theta < \theta_u^j) = \mathbb{E}(D^{(1)} + D^{(2)} | W = \varpi_j). \tag{13}$$

For any given $\tau \in [0, p_1] \subset [0, p_2]$. The one-to-one mappings from the type to the optimal effort and from the type to the realized cost, both conditional on W, imply the corresponding quantiles of $e_j(\cdot)$ and $c_j(\cdot)$ as

$$e_j(\theta(\tau)) = E_j(\tau/p_j), \quad c_j(\theta(\tau)) = C_j(\tau/p_j), \text{ for } j = 1, 2. \tag{14}$$

Consequently, a quantile version of (12) can be written as

$$\theta(\tau) = H^{-1}(C_1(\tau/p_1)) + E_1(\tau/p_1) = H^{-1}(C_2(\tau/p_2)) + E_2(\tau/p_2),$$

or equivalently

$$H^{-1}(C_1(\tau/p_1)) = H^{-1}(C_2(\tau/p_2)) + \Delta\tilde{E}(\tau), \tag{15}$$

where $\Delta\tilde{E}(\tau) \equiv E_2(\tau/p_2) - E_1(\tau/p_1)$. As shown in the proof of Proposition 3, $\Delta\tilde{E}(\tau) > 0$. Based on (15), we can establish the nonparametric identification of $H(\cdot)$. Specifically, for any $t \in [C_1(0), C_1(1)]$, it can be shown that

$$H^{-1}(t) = \begin{cases} \underline{\theta} - \underline{e}, & \text{for } t = C_1(0) \\ \underline{\theta} - \underline{e} + \sum_{k=0}^{\infty} \Delta\tilde{E}(\tau_k(t) \cdot p_1), & \text{for } t \in (C_1(0), C_1(1)] \end{cases} \tag{16}$$

where $\{\tau_k(t)\}_{k=0}^{\infty}$ is a unique sequence of quantiles. The sequence is identifiable since the distributions of C_1 and C_2 are identifiable. For a given τ_k, the term $\Delta\tilde{E}(\tau_k \cdot p_1) = e_2(\tau_k.p_1/p_2) - e_1(\tau_k)$ is also identifiable according to Proposition 2.

The unique sequence $\{\tau_k(t)\}_{k=0}^{\infty}$ is constructed as follows. Due to the the continuity and the strictly increase of cost function, for any $t \in (C_1(0), C_1(1)]$ there exists a unique $\tau_0(t) \in (0, 1]$ s.t. $C_1(\tau_0(t)) = t$. This implies that $t = C_1(\tau_0(t)) > C_2(\tau_0(t).p_1/p_2)$. Similarly, there exists a unique $\tau_1(t) \in (0, \tau_0(t))$ such that $C_1(\tau_1(t)) = C_2(\tau_0(t).p_1/p_2)$ and $C_1(\tau_1(t)) > C_2(\tau_1(t).p_1/p_2)$. Continuing such a procedure leads to a unique and strictly decreasing sequence $\{\tau_k(t)\}_{k=0}^{\infty}$ that satisfies the nonlinear recursive relation $C_1(\tau_{k+1}(t)) = C_2(\tau_k(t).p_1/p_2)$ with the initial condition $C_1(\tau_0(t)) = t$. Given the sequence of the quantiles, our main identification equation (16) can be derived from (15) by iteration:

$$\begin{aligned} H^{-1}(t) &= H^{-1}(C_1(\tau_0(t))) \\ &= H^{-1}(C_2(\tau_0(t).p_1/p_2)) + \Delta\tilde{E}(\tau_0(t).p_1) \\ &= H^{-1}(C_1(\tau_1(t)) + \Delta\tilde{E}(\tau_0(t).p_1) \\ &= H^{-1}(C_2(\tau_1(t).p_1/p_2)) + \Delta\tilde{E}(\tau_1(t).p_1) + \Delta\tilde{E}(\tau_0(t).p_1) \\ &= \cdots \\ &= H^{-1}(C_2(\tau_m(t).p_1/p_2)) + \sum_{k=0}^{m} \Delta\tilde{E}(\tau_k(t).p_1). \end{aligned} \tag{17}$$

Rearranging terms yields

$$\sum\nolimits_{k=0}^{m} \Delta\tilde{E}(\tau_k(t).p_1) = H^{-1}(t) - H^{-1}(C_2(\tau_m(t).p_1/p_2)).$$

In the appendix, we show that $\{\tau_k(t)\}_{k=0}^{\infty}$ is decreasing sequence that converge to 0. Taking limit on both side of the equation above yields

$$\sum\nolimits_{k=0}^{\infty} \Delta\tilde{E}(\tau_k(t).p_1) = H^{-1}(t) - H^{-1}(C_2(0)),$$

where $C_2(0)$ is the lower bound of C_2, whose corresponding argument of $H(\cdot)$ is $\underline{\theta} - \underline{e}$. This allows us to identify $H^{-1}(\cdot)$ as

$$H^{-1}(t) = \underline{\theta} - \underline{e} + \sum\nolimits_{k=0}^{\infty} \Delta\tilde{E}(\tau_k(t).p_1).$$

We summarize the result of identification the following proposition.

Proposition 3 *Under Assumptions 1-3, the cost function $H(\cdot)$ is nonparametrically identified on $[\underline{\theta} - \underline{e}, \theta_u^1 - e_u^1]$. Equivalently, the inverse function $H^{-1}(\cdot)$ is nonparametrically identified on $[\underline{c}, c_u^1]$.*

To the best of our knowledge, study on identification of cost function in contract models is very limited. Perrigne and Vuong (2012) show that the cost function is not identifiable under a static monopoly contract setting. The proposition above complements the existing literature with a positive identification result under a dynamic cost-based contract setting. Our positive result is based on: (i) identification of effort distribution following from recent development in measurement error literature, and (ii) the existence of an exclusion variable W.

It is worth noting that nonparametric identification of $H(\cdot)$ on $(\theta_u^1 - e_u^1, \overline{\theta}]$ is impossible. This is because an agent with $\theta \in (\theta_u^1, \overline{\theta}]$ would choose only CR contract, which provides little information on $H(\cdot)$. All we can learn from CR contract is $c = H(\theta)$ and the corresponding $e^* = 0$, while θ is unobserved.

Our identification results would still hold if Assumption 3 is replaced by a single crossing condition: ψ_1' and ψ_2' cross exactly once at some point on the intersection set of their interval supports. Under such a condition, the identification of the cost function $H(\cdot)$ follows similar steps to the proof of Proposition 3. We summarize this alternative assumption as follows.

Assumption 4 *There exists some $e_c \in [\underline{e}_1, e_u^1]$ such that $\psi_1'(e_c) = \psi_2'(e_c)$, $\psi_1'(e) > \psi_2'(e)$ for $e > e_c$ and $\psi_1'(e) < \psi_2'(e)$ for $e < e_c$, where $\underline{e}_1 \equiv e_1(\underline{\theta})$.*

Corollary 1 *Under Assumptions 1-2, 3(i) and 4, the cost function* $H(\cdot)$ *is identified on the interval* $[\max\{\underline{\theta} - e_j(\underline{\theta})\}, \min\{\theta_u^j - e_u^j\}], j = 1, 2$.

The proof of Corollary 1 is similar to that of Proposition 3. The main difference of identification in Corollary 1 is to first identify e_c, the intersection point of the two cost distributions under $W = \varpi_1$ and $W = \varpi_2$. Moreover, the identification result in Corollary 1 can be extended to the case with multiple crossing points because each intersecting point can be identified by using the identification argument in Corollary 1.

3.3 Identification of type distribution and other parameters

The identification of $H(\cdot)$ on part of its domain can be utilized to identify the distribution of innate cost θ on part of its support, which in turn is key to identify the ratio $\alpha/(1+\lambda)$, the disutility functions $\psi(\cdot) \equiv (\psi_1(\cdot), \psi_2(\cdot))'$, and agent's discount factor δ (or, equivalently, the intertemporal preference $r = 1/(1+\delta)$).

The first step is to recover θ from its relationship with the realized cost C. The basic idea is to use the one-to-one mapping between θ and $C = H(\theta - e(\theta))$. We have known by now that

$$\theta = \begin{cases} H^{-1}(c) + F_{E_1}^{-1}(F_{C_1}(c)), c \in [\underline{c}, c_u^1], \\ \\ H^{-1}(c), c \in (c_u^1, \overline{c}]. \end{cases} \tag{18}$$

where $F_{C_j}(\cdot)$, for $j = 1, 2$, is the distribution of cost among FP contracts, conditional on $W = \varpi_j$. And $F_{E_j}(\cdot)$, for $j = 1, 2$, is the distribution of effort among FP contracts, conditional on $W = \varpi_j$. Note that we can only recover the corresponding θ associated with $c \in [\underline{c}, c_u^1]$, which maps into $\theta \in [\underline{\theta}, \theta_u^1]$. This is because $H^{-1}(\cdot)$ is not identified for $c \in (c_u^1, \overline{c}]$. Consequently, we can identify the distribution of θ conditional on $\theta \in [\underline{\theta}, \theta_u^1]$.

Specifically, denote by $G(\cdot)$ and $g(\cdot)$ the truncated CDF and pdf of θ on $[\underline{\theta}, \theta_u^1]$. We can identify $G(\cdot)$ and $g(\cdot)$ according to recovered $\theta \in [\underline{\theta}, \theta_u^1]$. Similar identification approaches have been widely used to identify structural models, e.g., in Guerre et al. (2000) the distribution of bidders' valuations is recovered from observed bids using a similar method. Notice that $g(\cdot)$ is not the same as $f(\cdot)$, the full pdf of θ, on $[\underline{\theta}, \theta_u^1]$, but a rescaled version of it, as follows

$$g(\theta) = \begin{cases} \frac{f(\theta)}{F(\theta_u^1)}, \text{ if } \theta \in [\underline{\theta}, \theta_u^1], \\ \\ 0, \text{ otherwise}, \end{cases} \tag{19}$$

where $F(\theta_u^1) = \Pr(\underline{\theta} \leq \theta \leq \theta_u^1)$ is the probability that an agent chooses FF or CF contracts, and is identifiable via

$$F(\theta_u^1) = \mathbb{E}(D^{(1)} + D^{(2)}).$$

As a result, the pdf and CDF of θ is identified on $[\underline{\theta}, \theta_u^1]$ as

$$\begin{aligned} f(\theta) &= g(\theta)\mathbb{E}(D^{(1)} + D^{(2)}), \\ F(\theta) &= G(\theta)\mathbb{E}(D^{(1)} + D^{(2)}), \forall \theta \in [\underline{\theta}, \theta_u^1]. \end{aligned} \tag{20}$$

Next, we turn to the identification of the disutility functions of effort $\{\psi_1(\cdot), \psi_2(\cdot)\}$. Note that a direct implication from (10) and Proposition 2 is that, conditional on $W = \varpi_j$ for $j = 1, 2$, the optimal effort corresponding to any realized cost $c \in [\underline{c}, c_u^1]$ is identifible as

$$e_j^* = F_{E_j}^{-1}(F_{C_j}(c)), \;\; c \in [\underline{c}, c_u^1].$$

By combining this relationship with (18), we obtain a one-to-one mapping between the optimal effort e_j^* and agent's type θ, for all $\theta \in [\underline{\theta}, \theta_u^j]$, again conditional on $W = \varpi_j$. This mapping enables us to identify the derivative of disutility functions $\psi_j'(\cdot)$, $j = 1, 2$, from the first-order-condition of agent who exerts effort, i.e., $\psi_j'(e_j(\theta)) = H'(\theta - e_j(\theta))$ for all $\theta \in [\underline{\theta}, \theta_u^1] \subset [\underline{\theta}, \theta_u^2]$. An initial condition for this differential equation can be obtained from (7)

$$\psi_j(e_u^j) = \bar{b} - H(\theta_u^j - e_u^j) = \bar{b} - c_u^j.$$

Thus the solution for $\psi_j(\cdot)$ is

$$\psi_j(e) = \bar{b} - c_u^j - \int_e^{e_u^j} H'(e_j{}^{-1}(v) - v)dv, \;\; e \in [\underline{e}, e_u^j]. \tag{21}$$

Lastly, we focus on identifying the ratio $\alpha/(1+\lambda)$ which describes the relative weight the principal puts on agent's informational rent (profit) and social cost of public funds, and agent's discount factors δ, or equivalently intertemporal preference (weight) r with $r = 1/(1+\delta)$. Since the principal's optimization problem involves both types of agent which are induced to choose fixed-price and cost-reimbursement contracts, identification of the ratio requires information from all types of agent. We utilize the first-order-condition of the principal's problem (7) for identification. Note that $H'(\theta_u^1 - e_u^1)$, $f(\theta_u^1)$ and $F(\theta_u^1)$ on the right-hand-side are identified, it remains to to recover $F(\theta_l^1)$ and $H(\theta_u^1)$. First off,

$$F(\theta_l^1) = \Pr(\underline{\theta} \leq \theta \leq \theta_l^1) = \mathbb{E}(D^{(1)} | W = \varpi_1)$$

is just the probability that an agent chooses FF contract, thus $F(\theta_u^1) - F(\theta_l^1) = \mathbb{E}(D^{(2)} | \varpi_1)$. Our analysis of the theoretical model in Section 2 shows that an agent with innate cost θ_u^1

is indifferent between FP and CR contracts because both choices lead to zero profit. If an agent chooses the CR contract, $H(\theta_u^1) - \underline{c}_R = 0$, where $\underline{c}_R$ is the lower bound of the realized costs associated with CR contracts. Thus we can identify $H(\theta_u^1)$ as $\underline{c}_R$. Combining all the pieces above, we identify the ratio $\alpha/(1+\lambda)$ as

$$\frac{\alpha}{1+\lambda} = 1 - \frac{\underline{c}_R - \bar{b}}{H'(\theta_u^1 - e_u^1)} \frac{f(\theta_u^1)}{\mathbb{E}(D^{(2)}|\varpi_1)}. \tag{22}$$

Moreover, agent's discount factor δ is identified using the optimal condition (7)

$$\delta = \frac{\underline{b} - H(\theta_l^1 - e_l^1) - \psi_1(e_l^1)}{\bar{b} - \underline{b}}, \tag{23}$$

where θ_l^1 is identified as $\theta_l^1 = F^{-1}(\mathbb{E}(D^{(1)}|\varpi_1))$. The discount factor is a crucial objective to study agent's behavior and consequently conduct counterfactual or policy analyses. However, the discount factor oftentimes can not be identified in dynamic models. For example, Magnac and Thesmar (2002) show that decision makers' discount factor in dynamic discrete choice models can not be identified.

Theorem 1 *Let Assumptions 1-3 hold. Then the principal's relative ratio $\alpha/(1+\lambda)$ and agent's discount factor δ are identified. The distribution of agent's innate cost $F(\cdot)$, disutility function $\psi(\cdot) = (\psi_1(\cdot), \psi_2(\cdot))'$, and cost function $H(\cdot)$ are nonparametrically identified on $[\underline{\theta}, \theta_u^1]$, $[\underline{e}, e_u^1]$ and $[\underline{\theta} - \underline{e}, \theta_u^1 - e_u^1]$, respectively.*

The insight of identification in Theorem 1 naturally carries over to the one-period or static contract, since the equilibrium outcomes in the one-period setting is similar to that in the two-period contracts with renegotiation in the sense of providing similar information for identification. Thus the cost type distribution $F(\cdot)$, disutility function $\psi(\cdot)$ and cost function $H(\cdot)$ are nonparametrically identified on intervals corresponding to types associated with the choices of FP contracts in the static contract. Moreover, Theorem 1 also applies to the two-period contracts with commitment because the equilibrium outcome in the two-period contracts with commitment is just the twice-repeated version of that in the one-period (static) setting, as proved in Laffont and Tirole (1990).

Theorem 1 shows that the two-period FPCR model is nonparametrically identified for types associated with positive effort in at least one period. The results could be extended to a more general class of contracts: linear cost sharing-cost reimbursement (LCSCR) contracts where the fixed-price contract is replaced by a linear cost-sharing (LCS) contract. In a LCS contract, the payment is specified as a lump-sum payment q_0 plus a single fraction $\kappa \in [0,1]$ of realized cost for which the agent would be reimbursed.

$$q_t = q_0 + \kappa c_t, \kappa \in [0,1]. \tag{24}$$

Such a payment schedule is widely studied in the literature (e.g., Chu and Sappington, 2007) and nests the FPCR contract we analyzed as a special case with $\kappa = 0$ from the perspective of functional forms. Our procedure of identification could apply because linear cost sharing (LCS) contracts provide additional variation of payments by comparing with FP contracts where the payment is a constant and independent of cost.

Note that the upper bound of type's support, $\overline{\theta}$, is not required to obtain the results in Theorem 1. In actuality, the upper bound cannot be identified because it relies on the information of CR contracts. We show in the next section that parametrizing the cost function $H(\cdot)$ enables us to achieve identification of model primitives on their full support.

3.4 Semiparametric Identification

So far we have identified almost all model primitives nonparametrically, except $H(\cdot)$ on $(\theta_u^1 - e_u^1, \overline{\theta}]$, and $F_\theta(\cdot)$ on $(\theta_u^1, \overline{\theta}]$. Here we focus on the former, whose identification is the key to identify the latter. As previously discussed in the paper, $H(\cdot)$ is unidentified on $(\theta_u^1 - e_u^1, \overline{\theta}]$ due to a severe lack of identifying information from CC contracts. Our identification strategy for $H(\cdot)$ on $(\theta_u^1 - e_u^1, \overline{\theta}]$ is to develop additional regularity conditions which guarantee a unique extrapolation of $H(\cdot)$ from the already identified region $[\underline{\theta} - \underline{e}, \theta_u^1 - e_u^1]$ to $(\theta_u^1 - e_u^1, \overline{\theta}]$.

We regulate $H(\cdot)$ by parameterizing it. We assume that $H(\cdot)$ admits the parametric form $H(\cdot;\beta)$ on its entire domain $[\underline{\theta} - \underline{e}, \overline{\theta}]$, for a finite dimensional parameter $\beta \in \mathbb{R}^{K_\beta}$. Parameterizing the cost functions is a widely adopted procedure in the related literature. For example, Luo et al. (2018) parameterize the cost function to identify the truncated distribution of consumers' types. In our study, after parameterization, β is characterized by (16) as

$$H^{-1}(t;\beta) = \underline{\theta} - \underline{e} + \sum\nolimits_{k=0}^{\infty} \Delta \tilde{E}(\tau_k(t) \cdot p_1). \tag{25}$$

Note that the equation above holds for any given $t \in [\underline{\theta} - \underline{e}, \theta_u^1 - e_u^1]$, which provides strong identifying power. In many cases, if we parameterize $H(\cdot;\beta)$ carefully to avoid any trivial non-identification[6], then identification of β, or equivalently parametric identification of $H(\cdot)$ on $[\underline{\theta} - \underline{e}, \overline{\theta}]$, can be achieved.

Alternatively, we may identify $\beta \in \mathbb{R}^{K_\beta}$ via the key relationship (15), which, under

[6]Here, the term "trivial non-identification" refers to non-identification due to careless specification of parameters, which can be easily avoided. For example, if $H(\cdot)$ is parameterized as $H(t;\beta_1,\beta_2) \equiv (\beta_1 + \beta_2)[t - (\underline{\theta} - \underline{e})]^2 + \underline{c}$, then obviously (β_1, β_2) is unidentified. But the non-identification in this example can be easily avoid if we parameterize $H(\cdot)$ as $H(t;\beta) \equiv \beta[t - (\underline{\theta} - \underline{e})]^2 + \underline{c}$.

the parametric form, is as follows

$$H^{-1}(C_1(a_l);\beta) = H^{-1}(C_2(a_l \cdot \frac{p_1}{p_2});\beta) + E_2(a_l \cdot \frac{p_1}{p_2}) - E_1(a_l), \tag{26}$$

for a preselected sequence of constants $\{a_1, a_2, ..., a_L\} \subset (0,1)$, together with the initial condition $H^{-1}(\underline{c}) = \underline{\theta} - \underline{e}$. Roughly speaking, β is identified when L large enough. (Typically, $L \geq K_\beta$ is a minimum requirement, but dose not necessarily guarantee identification.) But detailed conditions for identification of nonlinear models (i.e. uniqueness of solution to a system of nonlinear equations, such as (26)) can be case dependent, and are generally complicated. Yet, for a special case when $H(\cdot)$ is specified to take a quadratic form, which we adopt in our empirical study, identification of β is straightforward based on (26). Note that once β is identified, we can recover the upper bound of type, $\overline{\theta}$ as $H^{-1}(\bar{c}_R;\beta)$ with $\bar{c}_R$ being the upper bound of realized costs among all CR contracts.

Once $H(\cdot)$ is identified on $[\underline{\theta} - \underline{e}, \overline{\theta}]$, identification of $F_\theta(\cdot)$ on its entire domains follows from (18). We summarize the identification results under parameterization of $H(\cdot)$ in the following theorem. Although $H(\cdot)$ is only parametrically identification, other unknown functions $F_\theta(\cdot)$ and $\psi(\cdot)$ are still nonparametrically identified, in the sense of which these can be viewed as semiparametric identification results.

Theorem 2 *Suppose Assumptions 1-3 hold and β is identified. Then $\alpha/(1+\lambda)$ and δ are identified. Moreover, $F(\cdot)$ and $\psi(\cdot)$ are nonparametrically identified on $[\underline{\theta}, \bar{\theta}]$ and $[\underline{e}, e_u^1]$, respectively.*

4 Estimation

In this section, we discuss estimation of the cost function $H(\cdot)$, the distribution of innate cost θ, the ratio $\alpha/(1+\lambda)$, and other model primitives. We provide a three-step semiparametric estimation scheme that follows our identification strategy closely. Firstly we estimate the distribution functions of the observed cost and latent effort level, respectively. Secondly we estimate $H(\cdot)$ as a plug-in estimator, relying on a sequence of quantiles estimated based on the estimated distribution functions from the first step. Finally, we estimate the distribution of θ and the ratio $\alpha/(1+\lambda)$. Here, we focus on the case of binary $W \in \{\varpi_1, \varpi_2\}$, which we prioritize throughout the paper.

Recall that the cost and disutility function, and the distribution of innate cost are all dependent on the covariates Z. To avoid "curse of dimensionality", it may be necessary to incorporate Z into our estimation parametrically instead of just conditioning on it. Nevertheless, modeling Z is application-specific and we will discuss the details in our empirical illustration. In this section, we will present our estimation procedure conditional on $Z = z$ for a given z for simplicity.

The sample $\{c_i, w_i, d_i^{(1)}, d_i^{(2)}, x_i, y_i\}_{i=1}^n$ is from observations on n two-period contracts, and is assumed to be i.i.d.. Among the n contracts, n_f of them are under a fixed price for at least one period (i.e. $d_i^{(1)} + d_i^{(2)} = 1$). Moreover, among the n_f fixed price contracts, $n_{f,1}$ of them are with $w_i = \varpi_1$, and $n_{f,2}$ of them are with $w_i = \varpi_2$. They form two subsamples $\{c_{1i}, d_{1i}^{(1)}, d_{1i}^{(2)}, x_{1i}, y_{1i}\}_{i=1}^{n_{f,1}}$ and $\{c_{2i}, d_{2i}^{(1)}, d_{2i}^{(2)}, x_{2i}, y_{2i}\}_{i=1}^{n_{f,2}}$, which turn out to be important for estimation.

Step 1: Estimation of Cost and Effort Distributions

The first step is to estimate the conditional distributions of cost C and optimal effort E^* given W.[7] For a binary W, this means to estimate the distributions of C and E^* with each of the two subsamples $\{c_{1i}, d_{1i}^{(1)}, d_{1i}^{(2)}, x_{1i}, y_{1i}\}_{i=1}^{n_{f,1}}$ and $\{c_{2i}, d_{2i}^{(1)}, d_{2i}^{(2)}, x_{2i}, y_{2i}\}_{i=1}^{n_{f,2}}$. Specifically, the density of E^* conditional on $W = \varpi_j$, denoted by f_{E_j} for $j = 1, 2$, is estimated along with some nuisance parameters by a sieve maximum likelihood estimator (SMLE) proposed in Shen (1997) as follows

$$(\widehat{m}, \widehat{f}_{E_j}, \widehat{f}_{V_1}, \widehat{f}_{V_2}) = \arg\max \sup_{(m, F_{E_j}, f_{V_1}, f_{V_2})} \sum_{i=1}^{n_{f,j}} \ln \int f_{V_1}(Y_{1i} - t) f_{V_2}(Y_{2i} - m(t)) f_{E_j}(t) dt, \tag{27}$$

where the max and sup are taken over suitably restricted sets of functions; $f_{V_1}(\cdot)$ and $f_{V_2}(\cdot)$, respectively, denote the densities of error terms V_1 and V_2. The optimization is subject to some restrictions which consist of constraints that the densities integrate to one and zero-mean constraints on the error densities $f_{V_1}(\cdot)$ and $f_{V_2}(\cdot)$. All unknown functions $f_{V_1}(\cdot)$, $f_{V_2}(\cdot)$, and $f_{E_j}(\cdot)$ are chosen in an appropriate sieve space constructed by truncated series such as Hermite orthogonal series with the number of terms in the series increasing with the sample size. More details about the implementation and properties of the sieve estimators can be found in Shen (1997), Ai and Chen (2003), Chen (2007), Carroll et al. (2010), and Chen et al. (2014). Depending on the parameters of interest, the unknown densities $f_{V_1}(\cdot)$, $f_{V_2}(\cdot)$, and $f_{E_j}(\cdot)$ can also be set semiparametrically or fully parametric.

Unlike the optimal effort E^*, the realized cost C is observed. Therefore, the CDF of the realized cost conditional on the exclusion variable ϖ_j, denoted by $F_{C_j}(\cdot)$ for $j = 1, 2$, can be more straightforwardly estimated by empirical distribution, kernel estimator, or some sieve estimator.

[7] Note that both the observable realized cost C and unobservable (optimal) effort E^* are dependent on the exclusion variable W.

Step 2: Estimation of the Cost Function

Due to the lack of nonparametric identification on its whole domain, we aim at parametric estimation of the cost function $H(\cdot)$. We note that non/semi-parametric estimation of $H(\cdot)$ on $[\underline{\theta} - \underline{e}, \theta_u^1 - e_u^1]$ is possible, following from Proposition 3. This might be of practical interest, and is left for future research.

$H(\cdot)$ is parametrized as $H(\cdot;\beta)$ for a finite dimensional parameter $\beta \in \mathbb{R}^{K_\beta}$. For a preselected sequence of grid points $\{a_l\}_{l=1}^L \subset (0,1)$, β is estimated as

$$\hat{\beta} = \underset{\beta: H(\underline{\theta}-\underline{e};\beta)=\underline{c}}{\operatorname{argmin}} \sum_{l=1}^{L} \left[H^{-1}(\widehat{C}_2(a_l \cdot \frac{\hat{p}_1}{\hat{p}_2}); \beta) - H^{-1}(\widehat{C}_1(a_l); \beta) + \widehat{E}_2(a_l \cdot \frac{\hat{p}_1}{\hat{p}_2}) - \widehat{E}_1(a_l) \right]^2, \quad (28)$$

where for $j = 1, 2$

$$\hat{p}_j = \frac{1}{n_{f,j}} \sum_{i=1}^{n_{f,j}} [d_{ji}^{(1)} + d_{ji}^{(2)}]. \quad (29)$$

$\widehat{C}_j(\tau)$ and $\widehat{E}_j(\tau)$ are the τ-th quantiles of C and E^* conditional on $W = \varpi_j$ and they are estimated according to $\widehat{F}_{C_j}$ and $\widehat{F}_{E_j}$ from the first step, respectively. $H^{-1}(\cdot;\beta)$ represents the inverse function of $H(\cdot)$ corresponding to the parametrization $H(\cdot;\beta)$. Consequently, $H(\cdot)$ and $H^{-1}(\cdot)$ are estimated as

$$\widehat{H}(\cdot) = H(\cdot; \hat{\beta}) \text{ and } \widehat{H}^{-1}(\cdot) = H^{-1}(\cdot; \hat{\beta}),$$

respectively.

For a special case, when $H(\cdot)$ is parametrized as a quadratic function, we can incorporate the initial condition $H(\underline{\theta} - \underline{e}) = \underline{c}$ directly into the parametrization, which yields

$$H(t;\beta) = \beta_2[t - (\underline{\theta} - \underline{e})]^2 + \beta_1[t - (\underline{\theta} - \underline{e})] + \underline{c}.$$

And it is easy to solve for the explicit form of $H^{-1}(\cdot;\beta)$ as

$$H^{-1}(c;\beta) = \underline{\theta} - \underline{e} + \frac{-\beta_1 + \sqrt{\beta_1^2 + 4\beta_2(c - \underline{c})}}{2\beta_2}.$$

In more general cases, an explicit form of the (real-valued) inverse function of $H(\cdot;\beta)$ may be very complicated, or may not even exist. For these cases, we suggest to parametrize $H^{-1}(\cdot)$ (rather than $H(\cdot)$) directly, and estimate it according to (28). After all, an estimated $H^{-1}(\cdot)$ is needed in the next step. Afterwards, $H(\cdot)$ can be estimated via certain extrapolation procedure based on $\widehat{H}^{-1}(\cdot)$.

Step 3: Estimation of θ, $\alpha/(1+\lambda)$, and Other Primitives

Based on the results from the two steps above, we can calculate a fitted value for θ_i as

$$\hat{\theta}_i = \begin{cases} \widehat{H}^{-1}(c_i) + \sum_{j=1,2} \mathbf{1}(w_i = \varpi_j) \cdot \widehat{F}_{E_j}^{-1}(\widehat{F}_{C_j}(c_i)), \text{ for FP contracts,} \\ \\ \widehat{H}^{-1}(c_i), \text{ for CR contracts.} \end{cases} \tag{30}$$

Then the CDF for θ is estimated as the empirical distribution based on the sample of fitted values $\{\hat{\theta}_i\}_{i=1}^n$,

$$\widehat{F}(t) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(\hat{\theta}_i \leq t). \tag{31}$$

And the pdf $f(\cdot)$ can be estimated as a kernel estimator calculated from $\{\hat{\theta}_i\}_{i=1}^n$.

Estimation of $\alpha/(1+\lambda)$ is based on (22)

$$\frac{\alpha}{1+\lambda} = 1 - \frac{\underline{c}_R - \bar{b}}{H'(\theta_u^1 - e_u^1)} \frac{f(\theta_u^1)}{\mathbb{E}(D^{(2)}|\varpi_1)}.$$

Specifically, since $F(\theta_u^1) = E(D^{(1)} + D^{(2)}|W = \varpi_1)$, θ_u^1 can be estimated as

$$\hat{\theta}_u^1 = \widehat{F}^{-1}\left(\frac{\sum_{i=1}^n \mathbf{1}(w_i = \varpi_1) \cdot (d_i^{(1)} + d_i^{(2)})}{\sum_{i=1}^n \mathbf{1}(w_i = \varpi_1)}\right). \tag{32}$$

Consequently $\alpha/(1+\lambda)$ is estimated as

$$\widehat{\alpha/(1+\lambda)} = 1 - \frac{\underline{c}_R - \bar{b}}{H'(\widehat{\theta}_u^1 - \hat{e}_u^1; \hat{\beta})} \frac{\hat{f}_\theta(\hat{\theta}_u^1) \cdot \sum_{i=1}^n \mathbf{1}(w_i = \varpi_1)}{\sum_{i=1}^n \mathbf{1}(w_i = \varpi_1) \cdot d_i^{(2)}}. \tag{33}$$

Estimation of $\psi_j(\cdot)$, for $j = 1, 2$, is based on the integration equation (21) and the fact that there is a one-to-one increasing mapping $e_j(\theta) = F_{E_j}^{-1}(F(\theta)/F(\theta_u^l)$ from $[\underline{\theta}, \theta_u^j]$ to $[\underline{e}, e_u^j]$. For any $e \in [\underline{e}, e_u^1]$, $\psi_j(e)$ is estimated as

$$\widehat{\psi}_j(e) = \bar{b} - c_u^j - \int_e^{\hat{e}_u^j} H'\left(\widehat{F}^{-1}(\widehat{F}_{E_j}(v) \cdot \widehat{F}(\theta_u^j)) - v; \hat{\beta}\right) dv, \tag{34}$$

where $\widehat{F}(\theta_u^j)$ can be calculated as

$$\widehat{F}(\theta_u^j) = \frac{\sum_{i=1}^n \mathbf{1}(w_i = \varpi_j) \cdot (d_i^{(1)} + d_i^{(2)})}{\sum_{i=1}^n \mathbf{1}(w_i = \varpi_j)}. \tag{35}$$

Finally, the discount factors δ_j, for $j = 1, 2$, are estimated based on (23), as follows

$$\hat{\delta} = \frac{\underline{b} - H(\hat{\theta}_l^1 - \hat{e}_l^1; \hat{\beta}) - \psi_1(\hat{e}_l^1)}{\bar{b} - \underline{b}}, \tag{36}$$

where

$$\hat{\theta}_l^1 = \widehat{F}^{-1}\left(\frac{\sum_{i=1}^n \mathbf{1}(w_i = \varpi_1) \cdot d_i^{(1)}}{\sum_{i=1}^n \mathbf{1}(w_i = \varpi_1)}\right) \tag{37}$$

and $\hat{e}_l^1 = \widehat{F}_{E_1}^{-1}\left(\widehat{F}_{C_1}(c_l^1)\right)$.

The estimators above are all consistent under regularity conditions. We formalize the consistency result in the following proposition.

Proposition 4 *Suppose Assumptions 1-3 hold. In addition, we assume: (i) $H(t; \beta)$ is continuous in both $t \in (\underline{\theta} - \underline{e}, \overline{\theta})$ and $\beta \in \mathcal{B}$, with $\mathcal{B}$ being compact, (ii) β is identified via (26), and (iii) $\{\hat{p}_1, \hat{p}_2, \widehat{C}_1(a_l), \widehat{C}_2(a_l \cdot \frac{p_1}{p_2}), \widehat{E}_1(a_l), \widehat{E}_2(a_l \cdot \frac{p_1}{p_2})\}_{l=1}^L$ are consistent. Then $\hat{\beta}$, $\{\hat{\theta}_i\}_{i=1}^n$, $\widehat{F}_\theta(\cdot)$, $\widehat{\alpha/(1+\lambda)}$, $\widehat{\psi}(\cdot)$, and $\hat{\delta}$ are all consistent.*

5 Empirical Illustration

In this section we apply our method to analyze transport procurement contracts. In France a local authority (principal) contracts with a single operator (agent) to provide the transport service by using fixed-price contracts or cost-reimbursement contracts. Gagnepain and Ivaldi (2002) confirmed through a test that adverse selection (on private innate cost of operators) and moral hazard (due to unobserved cost-reducing effort exerted by operators) are two important features of the industry. Regulatory rules require that these contracts must be renegotiated every five years between the two parties. Thus the dataset is particularly suitable for our model.

The main goal of our application is to evaluate the influence of different properties of cost-reducing effort on the social welfare. To do so, we estimate the structural parameters of the model by following the semiparametric estimation procedure proposed in Section 4, and then calculate the social welfare under monotone optimal effort and constant optimal effort, respectively. As shown earlier, the monotone optimal effort implied by convex cost functions means that the optimal effort is increasing in the innate cost, while the constant optimal effort implied by a linear cost function means that the optimal effort is independent of the innate cost. The latter property of optimal effort is widely involved in the related theoretical literature including Laffont and Tirole (1988, 1990), Rogerson (2003), Chu and Sappington (2007), Battaglini (2007), and Garrett (2014). The empirical evidence, however, suggests that the optimal effort is increasing in their innate costs, such as in the transport industry in France by Gagnepain and Ivaldi (2002) and electricity industry in the U.S. by Abito (2014). Hence, we aim to evaluate whether the social welfare of FPCR contracts under monotone optimal efforts significantly differs from that under constant optimal effort. This investigation is important both in theory and in practice.

Our empirical findings show that the cost function is significantly convex, thus implying the monotone optimal effort. The counterfactual results indicate that the difference of social welfare between monotone effort and constant effort is large, thus providing strong empirical evidence that (1) the theoretical literature should be cautious when specifying the cost function in simple contracts, and that (2) the performance of simple contracts varies with the property of optimal effort, which is determined by cost specifications.

5.1 Data

The dataset includes 543 two-period contracts implemented from 1987 to 2001. Among these contracts, 281 observations are two-period fixed-price contracts (FF), 88 observations are CR contract in the first period followed by one FP contract in the second period (CF), and the remaining 174 ones are a two-period CR contracts (CC). The dataset reports the type of contract, the realized cost for each contract, and the subsidy for each FP contract (i.e., the fixed prices paid from the principal to the agent). In addition, it provides some characteristics of operators/contracts, including the labor fee, the number of employees, the number of drivers, the size rolling stock (measured by the number of vehicles), and the ownership of operators.

Table 1 provides a summary statistics of the dataset. On average the cost is about 17 million euros and the subsidy is approximately 19 million per contract, which implies that on average the operators are profitable. The average labor fee is 10.7 million and accounts for 64 percent of the total cost, suggesting that reducing the labor fee is critical to increase the operator's profit. The average numbers of employees and drivers are 413 and 278, respectively. That is, on average more than one half of the employees are drivers, implying the intensive labor of transport industry. As the ownership indicates, one half of the operators is privately owned and another half is publicly owned.

Table 1: Summary Statistics

Variables	Mean	Std. Dev.	Min	Median	Max
# of Contracts	543				
# of FF	281				
# of CF	88				
# of CC	174				
Cost	16860	15954	2397	10347	93993
Subsidy	18794	18236	2265	12039	114483
Number of employees	413	364	68	267	1772
Number of drivers	278	216	47	144	1182
Labor fee	10740	10241	716	6650	53178
Rolling stock	165	121	33	84	724
Private owenship	0.52	0.50	0.00	1.00	1.00

All variables are real terms. The units of cost, subsidy and labor fee are 1000 euros.

5.2 Empirical strategies

The estimation strategy follows our semiparametric estimation procedure by choosing the two measurements (X, Y) and the exclusion variable W. One plausible variable for the measurement of effort X is the share of drivers among all the employees (employees consist of drivers and engineers) because the share of engineers provides a measure for the endowment of skills embodied in the operator. As Gagnepain et al. (2013) argue, engineers are generally responsible for research and development, quality control, maintenance, and efficiency. Since the optimal effort is positively related to the innate cost, we expect that the share of drivers in the total labor force to be positively related to the optimal effort. A natural choice for the second measurement of Y is the labor fee because labor fee represents 64 percent of the total cost, which can be interpreted as a function of the optimal effort due to the one-to-one mapping between the optimal effort and the innate cost. As Cicala (2015) suggests, one can employ cost-related variables to infer agent's effort.

Our choice of W is a dummy variable indicating the ownership of the operator, i.e., whether the operator is privately owned or not. Among the four operators in the dataset, three operators are privately owned ($W = \varpi_1$) while only Transdev is not privately owned ($W = \varpi_2$). As the empirical evidence in Gagnepain et al. (2013) shows, Transdev enjoys a less costly effort technology than other three private operators, which could be related to the operator's internal structure. Thus, we expect that Transdev takes a different disutility function of effort from other three operators. In addition, the covariate Z is taken as the number of vehicle in the operator's rolling stock. Recall that the operator's bargaining power α and the cost of public funds λ are not separately identified since only the ratio $\alpha/(1+\lambda)$ appears at the equilibrium conditions in (7). The empirical studies suggest that λ is in the interval $[0.15, 0.40]$ in an efficient tax systems (Ballard et al., 1985). We choose $\lambda = 0.3$ as in Gagnepain et al. (2013).

Because of the relative small sample size, we deviate slightly from the procedure in Section 4 to estimate model primitives. Specifically, we parametrize both the cost and the disutility function,

$$\begin{aligned} H(\theta - e) &= \beta_1(\theta - e) + \beta_2(\theta - e)^2, \\ \psi_j(e) &= \gamma_1^j e + \gamma_2^j e^2, j = 1, 2. \end{aligned} \tag{38}$$

The dependence of θ on the covariate Z is modeled as

$$\theta = \theta_0 + \lambda_\theta Z,$$

where θ_0 is a random variable. This specification above implies an additively separable form of the optimal effort (we drop the subscript j for ease of exposition)

$$e = e_0 + \lambda_e Z,$$

where e_0 and λ_e can be expressed as function of parameters of $H(\cdot), \psi_j(\cdot)$ and $\theta(\cdot)$.

$$e_0 = e_0(\theta_0; \beta_1, \beta_2, \lambda_1, \lambda_2); \;\; \lambda_e = \lambda_e(\beta_1, \beta_2, \lambda_1, \lambda_2, \lambda_\theta). \tag{39}$$

The decomposition above allows us to write the distribution of e conditional on Z as $f_{E|z}(e|z) = f_{E_0}(e - \lambda_e z)$, with $f_{E_0}(\cdot)$ being the density of E_0. Furthermore, the function $m(\cdot)$ is parametrized as follows.

$$m(E; \zeta) = \zeta_1 E + \zeta_2 E^2. \tag{40}$$

In the first step of estimation, where $\beta_1, \beta_2, \zeta_1$, and ζ_2 are estimated using (28), we use the Hermite orthogonal series $q_n(x)$ as our sieve basis functions.

$$q_n(x) = \sqrt{\frac{1}{\sqrt{\pi} n\,!2^n}} H_n(x) e^{-\frac{x^2}{2}},$$

where $H_0(x) = 1$, $H_1(x) = 2x$, and $H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x)$. The densities f_{V_1}, f_{V_2} and $f_{E_0^j}, j = 1, 2$ are expended by the sieve basis as:

$$f_{E_0^j}(x) = \Big(\sum\nolimits_{i=0}^{k_j} \rho_i q_i(x)\Big)^2, \; f_{V_1}(x) = \Big(\sum\nolimits_{i=0}^{m_1} \delta_i q_i(x)\Big)^2, \; f_{V_2}(x) = \Big(\sum\nolimits_{i=0}^{m_2} \pi_i q_i(x)\Big)^2,$$

where $k_j, m_j, j = 1, 2$ are smoothing parameters. In our estimation, we choose $k_1 = k_2 = m_1 = m_2 = 3$.

The objective $p_j, j = 1, 2$ also depends on Z and we estimate it using the standard kernel estimation. The parametrized disutility function $\psi_j(\cdot)$ is estimated by a minimum distance estimator based on the first-order-condition $H'(\theta - e_j) = \psi'_j(e_j)$, which is different from Section 4. The parameters α and δ (or equivalently r) are estimated using (33) and (36), respectively.

5.3 Estimation results

In Figure 1, we present estimates of densities $f_{E_0^j}(\cdot)$, together with their point-wise confidence intervals.[8] The estimates illustrate the distribution of optimal effort differs slightly across two types of firms. The average optimal effort for public firms is smaller that for private ones, and this result implies that public firms have less incentives. This observation is consistent with the existence of the exclusion variable W: some firms have larger marginal disutility than others and as a result the former types of firm exert less effort.

We present our estimation results in Table 2. The result indicates that the second measurement of optimal effort, i.e., the function $m(\cdot)$ is significantly convex in the optimal effort for the privately owned operators. The cost function $H(\cdot)$ is significantly convex in the operator's innate cost, thus providing strong empirical evidence against the linear cost function assumed in the related literature. An important implication of the convex cost function is that operators enjoy increasing returns to scale in the cost it induces, which provides less efficient operators with incentives to exert more effort. Furthermore, the convexity of cost is consistent with the convexity of the second measurement of the optimal effort for the privately owned operators since the labor fee, which is the the second measurement of the optimal effort, accounts for more than one half of the total cost.

The disutility functions $\psi_j(\cdot)$ are both significantly linear, thus implying that the marginal disutility of exerting cost-reducing effort is constant across different levels of effort. Focusing only on γ_1, the estimates are consistent with our assumption that the marginal disutility for public firms ($W = \varpi_2$) is larger. The innate cost is decreasing in the operator's rolling stock (captured by the parameter λ_θ), which suggests that the efficiency (productivity) of the operator is increasing in the operator's size represented by the rolling stock. The intertemporal weight r is 0.038, impling that operators pay most attention to the profit of the second-period. The operator's bargaining power is 1.299, as

[8] Recall that $e = e_0 + \lambda_e Z$. e_0 can be negative even though we require $e \geq 0$.

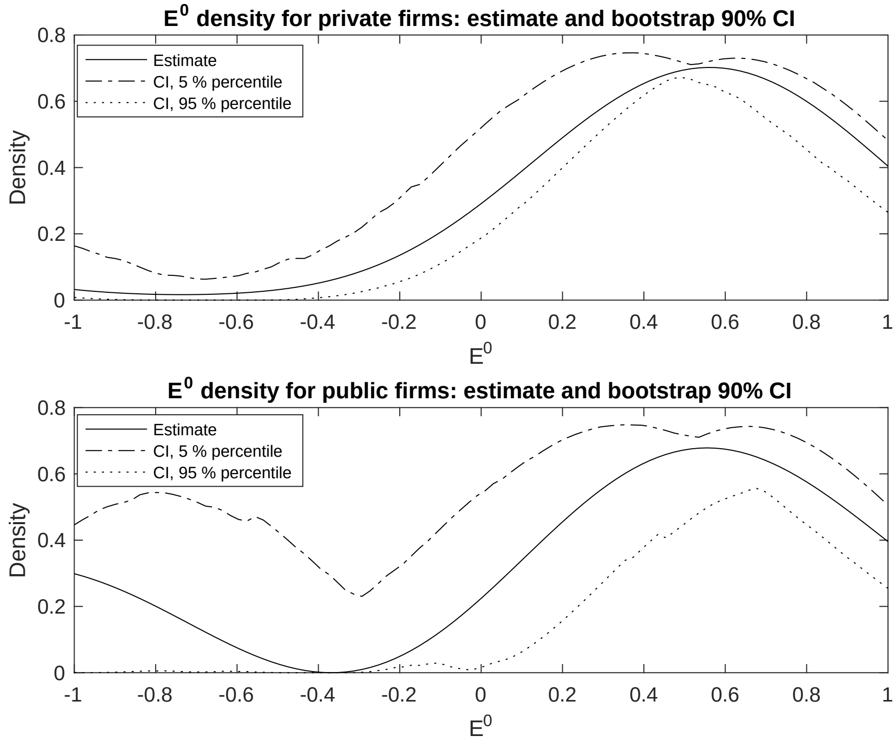
Table 2: Estimation results of model parameters

Functions/Parameters	Variables	Private($W = \varpi_1$)	Public ($W = \varpi_2$)
Measurement Y	Effort (ζ_1)	11.374***	7.493
		(3.870)	(6.417)
	Effort × Effort (ζ_2)	11.879***	7.740
		(3.218)	(8.253)
Cost	Ineff. (β_1)	0.430	
		(64.706)	
	Ineff. × Ineff. (β_2)	25832.4*	
		(13969.0)	
Disutility	Effort (γ_1)	35909.0***	40063.3***
		(12787.0)	(13815.4)
	Effort × Effort (γ_2)	1778.441	1682.441
		(1261.966)	(8113.650)
Innate cost	Rolling stock (λ_θ)	-0.0004***	
		(0.0001)	
Intertemporal weight (r)		0.038	
		(0.174)	
Bargaining power (α)		1.299***	
		(0.026)	
Sample size		369	

Standard errors in parentheses are bootstrapped 1000 times. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

We estimate parameters λ_θ, r and α conditional on $W = \varpi_1$ and $W = \varpi_2$ separately, then take the average.

Figure 1: Densities of optimal effort



is required by the theoretical restriction that $\alpha < 1 + \lambda$.[9]

5.4 Counterfactuals

In this section we assess the social welfare that would be achieved if the optimal effort were constant implied by a linear cost specification, which is widely assumed in the literature, and make comparisons of social welfare between constant optimal effort and monotone optimal effort.

The cost function is often specified as a linear function in many existing literature,

$$H(\theta - e) = \beta(\theta - e).$$

A consequence of the specification above is that the optimal effort is independent of the innate cost θ. This is implied by the first order condition

$$H' = \beta = \psi'(e^*).$$

[9]The parameters λ_θ, r and α are first estimated based on $W = \varpi_1$ and $W = \varpi_2$ separately, then we take the average as our estimates.

Table 3: Counterfactual results

Parameters	Cost	Intertemporal weight	Bargaining power
Est.	23465.4***	0.506***	0.571
	(5177.8)	(0.115)	(0.388)

Standard errors in parentheses are bootstrapped 1000 times.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

That is, the optimal effort is constant for any type of the agent associated with FP contracts. Since the optimal effort has no variation, the disutility function $\psi(\cdot)$ can not be identified. Therefore, we normalize the disutility function to be

$$\psi(e) = e^2,$$

which satisfies Assumption 1.[10] The details on semiparametric identification and constructive estimation of the model with constant optimal effort are provided in the Appendix.

The estimation results under the specification of linear cost function are presented in Table 3. As the estimation results show, the cost is significantly linear in innate cost, which is different from the prior empirical result where the cost is not significant in the linear term of innate cost. Given the significant convexity of cost in Table 2, the widely used linear cost function bears high risk of mis-specification, which may result in misleading conclusions. The intertemporal weight is 0.506, which suggests that the operator pays equal attention to the profit of both periods, while Table 2 implies that the operator pays most attention to the profit of the second period. The bargaining power is 0.517, which is only one half of the bargaining power in Table 2. Now we calculate and compare the social welfare under both monotone optimal effort and constant optimal effort. Let SW^M and SW^C denote the social welfare of the contracts under monotone optimal effort and constant optimal effort, respectively.[11]

$$\begin{aligned} SW^M(z) &= S - (1+\lambda)T^M(z) + \alpha^M U^M(z), \\ SW^C(z) &= S - (1+\lambda)T^C(z) + \alpha^C U^C(z), \end{aligned} \tag{41}$$

where $T^M(z)$ and $T^C(z)$ are subsidy (tax) under monotone optimal effort and constant optimal effort, respectively, $U^M(z)$ and $U^C(z)$ are their informational rent (profit) counterparts, and α^M and α^C are their respective bargaining powers.

[10]We also conduct counterfactuals with different normalizations of $\psi(\cdot)$ by choosing different values of $\chi > 0$ in the general form $\psi(e) = \chi e^2$, which is widely used in the related theoretical literature such as Laffont and Tirole (1988), Rogerson (2003), Chu and Sappington (2007), and Battaglini (2007).

[11]The social welfare SW^M depends on the exclusion variable $W \in \{\varpi_1, \varpi_2\}$. We take the average of the social welfare under ϖ_1 and ϖ_2 in the counterfactual analysis.

The definition of these objectives are as follows. The tax under monotone optimal efforts is defined as

$$T^M(z) = \underline{b}F_{\theta|z}(\theta_1^*(z)|z) + \int_{\theta_1^*(z)}^{\theta_2^*(z)} [r^M H(\theta) + (1-r^M)\bar{b}]dF_{\theta|z}(\theta|z) + \int_{\theta_2^*(z)}^{\infty} H(\theta)dF_{\theta|z}(\theta|z),$$

and $T^C(z)$ is defined similarly. The utility under monotone optimal efforts is defined as

$$\begin{aligned} U^M(z) &= \int_{-\infty}^{\theta_1^*(z)} [\underline{b} - H(\theta - e(\theta)) - \psi(e(\theta))]dF_{\theta|z}(\theta|z) \\ &+ (1-r^M)\int_{\theta_1^*(z)}^{\theta_2^*(z)} [\bar{b} - H(\theta - e(\theta)) - \psi(e(\theta))]dF_{\theta|z}(\theta|z), \end{aligned}$$

and $U^C(z)$ is defined similarly except that the optimal effort $e(\theta)$ is replaced by the constant optimal effort e^*. Consequently, the welfare difference is

$$\Delta SW(z) \equiv SW^C(z) - SW^M(z) = \big[\alpha^C U^C(z) - \alpha^M U^M(z)\big] - (1+\lambda)\big[T^C(z) - T^M(z)\big],$$

where the first term on the right side is the difference of bargaining power-weighted informational rent, and the second term is the difference of social cost. By integrating out the covariates z, we obtain the average social welfare difference $\overline{\Delta SW} = \int_{z\in\mathcal{Z}} \Delta SW(z)dG(z)$ by plugging their corresponding estimates of parameters into their definitions, where $G(\cdot)$ is the cdf of Z.[12]

Table 4: Welfare Differentials for the average network

Welfare items	Model	Estimate
Social cost	Monotone $(1+\lambda)T^M$	23.4
	Constant $(1+\lambda)T^C$	25.1
	Differential $(1+\lambda)(T^C - T^M)$	1.7
Weighted informational rent	Monotone $\alpha^M U^M$	23.2
	Constant $\alpha^C U^C$	10.3
	Differential $\alpha^M U^M - \alpha^C U^C$	12.9
Welfare differential ΔSW	$(\alpha^M U^M - \alpha^C U^C) - (1+\lambda)(T^C - T^M)$	14.6

All estimates are in million euros.

[12] We restrict the sample to FP contracts (including FF choices and CF choices) because a fixed-price is unobserved for CC contracts.

Based on Table 4, the difference of social welfare between monotone effort and constant effort is 14.6 million euros. The large difference is mainly due to the difference of weighted informational rent 12.9 million euros (56 percent), and the difference of social cost is only 1.7 million euros (7 percent). The substantial difference of social welfare is reminiscent of An and Zhang (2018) who prove that in a static (one-period) FPCR contract, the social welfare of FPCR contract under monotone optimal efforts may differ substantially from that under constant optimal effort, and the magnitude of difference relies crucially on the magnitude of monotonicity of the optimal effort. Intuitively, under the linear cost function an operator has no incentive to exert more effort than others regardless of what its innate cost is. Nevertheless, when the cost function is convex, an operator with a higher innate cost enjoys a larger reduction of cost than a lower-innate cost operator by exerting the same effort. This may explain the result that the weighted informational rent under monotone effort is much larger than that under constant effort. The smaller difference of social cost is due to the comparability of the estimates of cost functions in terms of inducing similar costs.

These findings suggest that it is crucial to take into account the monotonicity of optimal effort, which is determined by the functional form of agent's cost function, when we evaluate the performance of FPCR menus.

6 Conclusion

We provided a rigorous econometric framework to analyze two-period FPCR contracts with renegotiation. We proved that the model is nonparametrically identified on intervals corresponding to FF and CF contracts. Further we provided semi-nonparametric identification results on intervals corresponding to CC contracts. Our identification results are applicable to a large class of simple contracts. Based on the identification strategy, we proposed a semiparametric procedure to estimate the model primitives. In the empirical study, using data from public transport procurement contracts in France, we found that cost function of operators are significantly convex, thus providing empirical evidence that the optimal effort is increasing in the operator's type, rather than being a constant as implied by a linear specification of the cost function, which is widely adopted in the related literature. In addition, the monotonicity of optimal effort has important implications for the welfare analysis: the social welfare is much smaller under a constant optimal effort than that under monotone optimal effort.

References

Abito, J. M., "Agency Costs in Environmental Regulation: Evidence from Regulated Electric Utilities," Working Paper, University of Pennsylvania, 2014. 5

Ai, C. and X. Chen, "Efficient estimation of models with conditional moment restrictions containing unknown functions," *Econometrica* 71 (2003), 1795–1843. 4

An, Y. and D. Zhang, "Simple Menus of Cost-based Contracts with Monotone Optimal Efforts," *forthcoming, Journal of Public Economic Theory* (2018). 5.4

Aryal, G., I. Perrigne and Q. Vuong, "Identification of Insurance Models with Multidimensional Screening," Working paper, Australian National University, 2010. 1

Bajari, P. and S. Tadelis, "Incentives versus transaction costs: A theory of procurement contracts," *RAND Journal of Economics* 32 (2001), 387–407. 1

Ballard, C. L., J. B. Shoven and J. Whalley, "General equilibrium computations of the marginal welfare costs of taxes in the United States," *American Economic Review* 75 (1985), 128–138. 5.2

Banerjee, A. V. and E. Duflo, "Reputation Effects and the Limits of Contracting: A Study of the Indian Software Industry," *Quarterly Journal of Economics* 115 (2000), 989–1017. 1

Baron and Myerson, "Regulating a Monopolist with Unknown Costs," *Econometrica* 50 (1982), 911–930. 2.1

Baron, D. P., "Regulation and legislative choice," *RAND Journal of Economics* 19 (1988), 467–477. 2.1

Battaglini, M., "Optimality and renegotiation in dynamic contracting," *Games and economic behavior* 60 (2007), 213–246. 5, 10

Carroll, R. J., X. Chen and Y. Hu, "Identification and estimation of nonlinear models using two samples with nonclassical measurement errors," *Journal of nonparametric statistics* 22 (2010), 379–399. 4

Chen, X., "Large sample sieve estimation of semi-nonparametric models," *Handbook of econometrics* 6 (2007), 5549–5632. 4

Chen, X., Z. Liao and Y. Sun, “Sieve inference on possibly misspecified semi-nonparametric time series models,” *Journal of Econometrics* 178 (2014), 639–658. 4

Chernozhukov, V. and C. Hansen, “An IV model of quantile treatment effects,” *Econometrica* 73 (2005), 245–261. 3.2

Chesher, A., “Identification in nonseparable models,” *Econometrica* 71 (2003), 1405–1441. 3.2

Chu, L. Y. and D. E. Sappington, “Simple cost-sharing contracts,” *American Economic Review* 97 (2007), 419–428. 1, 3.3, 5, 10

Cicala, S., “When Does Regulation Distort Costs? Lessons from Fuel Procurement in US Electricity Generation,” *American Economic Review* 105 (2015), 411–444. 3.1, 5.2

D’Haultfoeuille, X. and P. Février, “The provision of wage incentives: A structural estimation using contracts variation,” Working paper, CREST-ENSAE, 2015. 1

Faure-Grimaud, A. and D. Martimort, “Regulatory inertia,” *Rand Journal of Economics* 34 (2003), 413–437. 2.1

Gagnepain, P. and M. Ivaldi, “Incentive regulatory policies: the case of public transit systems in France,” *RAND Journal of Economics* 33 (2002), 605–629. 5

Gagnepain, P., M. Ivaldi and D. Martimort, “The Cost of Contract Renegotiation: Evidence from the Local Public Sector,” *American Economic Review* 103 (2013), 2352–83. 1, 3.2, 5.2, A

Garrett, D. F., “Robustness of simple menus of contracts in cost-based procurement,” *Games and Economic Behavior* 87 (2014), 631–641. 5

Gopal, A., K. Sivaramakrishnan, M. S. Krishnan and T. Mukhopadhyay, “Contracts in offshore software development: An empirical analysis,” *Management Science* 49 (2003), 1671–1683. 1

Guerre, E., I. Perrigne and Q. Vuong, “Optimal Nonparametric Estimation of First-Price Auctions,” *Econometrica* 68 (2000), 525–574. 3.3

Heckman, J. J., R. L. Matzkin and L. Nesheim, “Nonparametric identification and estimation of nonadditive hedonic models,” *Econometrica* 78 (2010), 1569–1591. 3.2

Hu, Y., "The Econometrics of Unobservables: Applications of Measurement Error Models in Empirical Industrial Rrganization and Labor Economics," *Journal of Econometrics* 200 (2017), 154–168. 1

Laffont, J.-J. and J. Tirole, "Using cost observation to regulate firms," *Journal of Political Economy* 94 (1986), 614–641. 1

———, "The dynamics of incentive contracts," *Econometrica* (1988), 1153–1175. 1, 5, 10

———, "Adverse selection and renegotiation in procurement," *Review of Economic Studies* 57 (1990), 597–625. 1, 2, 3.3, 5

———, *A Theory of Incentives in Procurement and Regulation* (MIT Press, 1993). 2.2

Li, T., "Robust and consistent estimation of nonlinear errors-in-variables models," *Journal of Econometrics* 110 (2002), 1–26. 3.1

Luo, Y., I. Perrigne and Q. Vuong, "Structural Analysis of Nonlinear Pricing," *forthcoming, Journal of Political Economy* (2018). 1, 3.4

Magnac, T. and D. Thesmar, "Identifying dynamic discrete decision processes," *Econometrica* 70 (2002), 801–816. 3.3

Perrigne, I. and Q. Vuong, "Nonparametric identification of a contract model with adverse selection and moral hazard," *Econometrica* 79 (2011), 1499–1539. 1, 5, 3.1

———, "On the identification of the procurement model," *Economics Letters* 114 (2012), 9–11. 5, 3.2

Rogerson, W. P., "An economic framework for analyzing DoD profit policy," Technical Report, RAND CORP SANTA MONICA CA, 1992. 1

———, "Simple menus of contracts in cost-based procurement and regulation," *American Economic Review* 93 (2003), 919–926. 1, 5, 10

Schennach, S. M. and Y. Hu, "Nonparametric identification and semiparametric estimation of classical measurement error models without side information," *Journal of the American Statistical Association* 108 (2013), 177–186. 1, 3.1, 3.1

Shen, X., "On methods of sieves and penalization," *The Annals of Statistics* (1997), 2555–2591. 4, 4

TORGOVITSKY, A., "Identification of nonseparable models using instruments with small support," *Econometrica* 83 (2015), 1185–1197. 3.2

VAN DER VAART, A., *Asymptotic Statistics* (Cambridge University Press, 1999). A

VIVES, X., *Oligopoly pricing: old ideas and new tools* (MIT press, 2001). 13

Appendix

A Proofs

Proof of Lemma 1. For a cost type associated with the FP contract, the assumption $H''(\cdot) > 0$ implies that the agent's objective function

$$\pi(\theta, e) \equiv q - H(\theta - e) - \psi(e) \tag{A.1}$$

is supermodular in θ and e, and hence the optimal effort e^* is (weakly) increasing in type θ by Topkis's theorem.[13] Under the additional assumption $\psi''(\cdot) > 0$, the optimal effort is single-valued due to the fact that the agent's objective function $\pi(\theta, e)$ is strictly concave in e. Since $\pi(\theta, e)$ is continuously differentiable in e, we obtain that the optimal effort is strictly increasing in type, that is, $e^{*'}(\theta) > 0$. Therefore, a less efficient agent will exert more effort, because an agent with a higher cost type enjoys a larger cost reduction than a lower cost type whenever exerting the same level of effort. Furthermore, the first-order-condition of (A.1) with respect to effort e is

$$H'(\theta - e^*) = \psi'(e^*). \tag{A.2}$$

The derivative of the equation above with respect to θ on both sides leads to

$$\psi''(e^*(\theta))e^{*'}(\theta) = H''(\theta - e^*(\theta))(1 - e^{*'}(\theta)). \tag{A.3}$$

Under Assumption 1, (A.3) implies that $0 < e^{*'}(\theta) < 1$ for any type θ with the FP contract.

Proof of Proposition 1. The proof below is similar to the the arguments in Gagnepain et al. (2013). Let $\Theta_G \equiv [\underline{\theta}, \theta_l]$, $\Theta_I \equiv (\theta_l, \theta_u]$, and $\Theta_B \equiv (\theta_u, \bar{\theta}]$. Denote respectively by

[13] See Theorem 2.3 in Vives (2001).

$C_1^0 = (b_1, b_2^0)$, $C_2^0 = (H(\theta), b_3^0)$ and $C_3^0 = (H(\theta), H(\theta))$ the agent's payments under each scenario Θ_G, Θ_I, Θ_B, and by $C^0 = (b_1, b_2^0, b_3^0)$ the overall menu of fixed prices. Denote $\widetilde{R} = (\widetilde{C}_2, \widetilde{C}_3) = (\widetilde{b}_2, \widetilde{b}_3)$ as a subsidy profile offered at the renegotiation stage following an initial offer C^0 and

$$\tilde{b}_2 \geq b_2^0 \ \text{ and } \ \tilde{b}_3 \geq b_3^0. \tag{A.4}$$

For simplicity, we decompose the proof into two lemmas.

Lemma A.1 ***(Renegotiation-proof)*** *There is no loss of generality in restricting the analysis to contracts of the form $C = (b_1, R)$ that come unchanged through the renegotiation process, such that $R = (b_2, b_3)$ maximizes the principal's second period welfare subject to the following acceptance conditions:*

$$\tilde{b}_2 \geq b_2 \ \textit{ and } \ \tilde{b}_3 \geq b_3. \tag{A.5}$$

Proof of Lemma A.1: For any initial contract C^0 and consider a renegotiated offer $\widetilde{R} = (\widetilde{b}_2, \widetilde{b}_3)$ that satisfies (A.4). Given the the agent's conjecture about the renegotiated offer $R = (b_2, b_3)$, the principal's expected welfare for date 2 becomes[14]

$$\begin{aligned} SW_2(C^0, \widetilde{R}, R) &= \int_{\underline{\theta}}^{\theta_l} \Big(S - (1+\lambda)\widetilde{b}_2 + \alpha(\widetilde{b}_2 - H(\theta - e^*(\theta)) - \psi(e^*(\theta)))\Big)\, dF(\theta) \\ &+ \int_{\theta_1^*}^{\theta_u} \Big(S - (1+\lambda)\widetilde{b}_3 + \alpha(\widetilde{b}_3 - H(\theta - e^*(\theta)) - \psi(e^*(\theta)))\Big)\, dF(\theta) \\ &+ \int_{\theta_2^*}^{\bar{\theta}} (S - (1+\lambda)H(\theta))\, f(\theta) d\theta \end{aligned} \tag{A.6}$$

Then the renegotiated offers $R = (b_2, b_3)$ must solve

$$R = \arg\max_{\widetilde{R}} SW_2(C^0, \widetilde{R}, R) \quad \text{subject to (A.4).} \tag{$\mathcal{R}^0$}$$

Due to the arbitrary C^0, it is easy to obtain that R also solves the following problem

$$R = \arg\max_{\widetilde{R}} SW_2(C \equiv (b_1, R), \widetilde{R}, R) \quad \text{subject to (A.5).} \tag{$\mathcal{R}$}$$

This completes the proof of Lemma A.1.

Let us now characterize renegotiation-proof allocations by solving the problem $\mathcal{R}$.

[14] Note that in $SW_2(C^0, \widetilde{R}, R)$, $\theta_l = \theta_l(b_1, b_2, b_3)$ and $\theta_u = \theta_u(\widetilde{b}_3)$ not $\theta_u(b_3)$.

Lemma A.2 *A first-period menu of contracts $C = (b_1, b_2, b_3)$ is renegotiation-proof if and only if the following two conditions hold:*

$$\theta_u(b_3) \geq \theta_l(b_1, b_2, b_3) \qquad (A.7)$$

$$\left(1 - \frac{\alpha}{1+\lambda}\right)[F(\theta_u) - F(\theta_l)] = \frac{[H(\theta_u) - b_3]f(\theta_u)}{H'(\theta_u - e^*(\theta_u))} \qquad (A.8)$$

Condition (A.7) guarantees that the interval Θ_I is non-empty.

Proof of Lemma A.2. First note that the assumption $\alpha < 1 + \lambda$ implies that the maximum of the integral in (A.6) requires that (A.5) is binding. Assume that $(\mathcal{R})$ is strictly quasi-concave in $\widetilde{b}_3$. The first-order condition of the optimization problem $(\mathcal{R})$ with respect to $\widetilde{b}_3$ at $\widetilde{b}_3 = b_3$ is

$$\begin{aligned} 0 &= \frac{d\theta_u}{db_3}\{S - (1+\lambda)b_3 + \alpha[b_3 - H(\theta_u - e^*(\theta_u)) - \psi(e^*(\theta_u))]\} f(\theta_u) \\ &\quad + \int_{\theta_1^*}^{\theta_u} (\alpha - 1 - \lambda) f(\theta) d\theta - \frac{d\theta_u}{db_3}[S - (1+\lambda)H(\theta_u)] f(\theta_u) \\ &= \frac{d\theta_u}{db_3}(1+\lambda)[H(\theta_u) - b_3] f(\theta_u) + \int_{\theta_1^*}^{\theta_u} (\alpha - 1 - \lambda) f(\theta) d\theta, \end{aligned}$$

where $b_3 = H(\theta_u - e^*(\theta_u)) + \psi(e^*(\theta_u))$ will be proved in (A.10) below. Note that $1 = [H'(\theta_u - e^*(\theta_u))(1 - e^{*'}(\theta_u)) + \psi'(e^*(\theta_u))e^{*'}(\theta_u)]d\theta_u/db_3 = H'(\theta_u - e^*(\theta_u))d\theta_u/db_3$, we obtain

$$(1 + \lambda - \alpha)[F(\theta_u) - F(\theta_l)] = \frac{(1+\lambda)[H(\theta_u) - b_3]f(\theta_u)}{H'(\theta_u - e^*(\theta_u))},$$

which completes the proof of Lemma A.2.

Define now the principal's intertemporal welfare when offering $C = (b_1, b_2, b_3)$ as[15]

$$\begin{aligned} SW(C) &= \int_{\underline{\theta}}^{\theta_l} \{S - (1+\lambda)(rb_1 + (1-r)b_2) + \alpha[rb_1 + (1-r)b_2 - H(\theta - e^*(\theta)) - \psi(e^*(\theta))]\} dF(\theta) \\ &+ \int_{\theta_1^*}^{\theta_u} \{S - (1+\lambda)(rH(\theta) + (1-r)b_3) + \alpha(1-r)[b_3 - H(\theta - e^*(\theta)) - \psi(e^*(\theta))]\} dF(\theta) \\ &+ \int_{\theta_2^*}^{\bar{\theta}} [S - (1+\lambda)H(\theta)] dF(\theta) \end{aligned}$$

The optimal renegotiation-proof menu solves the following optimization problem:[16]

$$\max_C SW(C) \quad \text{subject to (A.8).} \qquad (\mathcal{P}^R)$$

[15]Lemma (A.2) implies that $\theta_u = \theta_u(b_3)$ in $SW(C)$, and still $\theta_l = \theta_l(b_1, b_2, b_3)$.

[16]We assume (A.7) holds with strict inequality and (A.8) holds with equality as shown in Lemma A.2.

Then the two cut-off types under renegotiation satisfy

$$b_1 + \frac{(1-r)}{r}(b_2 - b_3) = H(\theta_l - e^*(\theta_l)) + \psi(e^*(\theta_l)), \tag{A.9}$$

$$b_3 = H(\theta_u - e(\theta_u)) + \psi(e(\theta_u)). \tag{A.10}$$

Due to the fact that $\frac{d\theta_l}{db_1} = \frac{r}{1-r}\frac{d\theta_l}{db_2}$ by (A.9), it is easy to show that the first-order conditions for b_1 and b_2 are the same, thus leading to the same optimal solution $b_1^R = b_2^R \equiv \underline{b}$. Denote the optimal solution for b_3 by $\bar{b}$. The first-order conditions with respect to b_1 and b_3 leads to

$$\frac{r(1+\lambda-\alpha)F(\theta_l)}{f(\theta_l)} = \frac{(1+\lambda)[rH(\theta_l) + (1-r)\bar{b} - \underline{b}] - \vartheta\left(1 - \frac{\alpha}{1+\lambda}\right)}{H'(\theta_l - e^*(\theta_l, w))},$$

$$\left(1 - \frac{\alpha}{1+\lambda}\right)[F(\theta_u) - F(\theta_l)] = \frac{[H(\theta_u) - \bar{b}]f(\theta_u)}{H'(\theta_u - e^*(\theta_u))} - \frac{[rH(\theta_l) + \bar{b}(1-r) - \underline{b}]f(\theta_l)}{rH'(\theta_l - e^*(\theta_l))}$$
$$+ \quad \frac{\vartheta m(\theta_l, \theta_u, \lambda, r, \alpha)}{(1+\lambda)(1-r)},$$

where $\vartheta > 0$ is the Lagrange multiplier of the renegotiation-proof constraint (A.8), and

$$m(\theta_l, \theta_u, \lambda, r, \alpha) = \left(1 - \frac{\alpha}{1+\lambda}\right)\left(\frac{f(\theta_u)}{H'(\theta_u - e^*(\theta_u))} - \frac{f(\theta_l)(r-1)}{rH'(\theta_l - e^*(\theta_l))}\right)$$
$$- \frac{[H'(\theta_u)f(\theta_u) + (H(\theta_u) - \bar{b})f'(\theta_u)]H'[\theta_u - e^*(\theta_u)]}{[H'(\theta_u - e^*(\theta_u))]^3}$$
$$+ \frac{H''[\theta_u - e^*(\theta_u)][1 - e^{*\prime}(\theta_u)][H(\theta_u) - \bar{b}]f(\theta_u)}{[H'(\theta_u - e^*(\theta_u))]^3}$$

This completes the proof of Proposition 1.

Proof of Lemma 2. Let $(\widetilde{C}, \widetilde{D}^{(1)}, \widetilde{D}^{(2)}, \widetilde{\overline{B}}, \widetilde{\underline{B}})$ denote the endogenous variables under the structure $\widetilde{\mathcal{S}}$. In actuality, the equivalence between $\mathcal{S}$ and $\widetilde{\mathcal{S}}$ can be obtained by taking a linear transformation that $\widetilde{\theta} = \xi_1\theta$ with $\xi_1 > 0$. To do this, let us first consider a general linear transformation that $\widetilde{\theta} = \xi_0 + \xi_1\theta$ with $(\xi_0, \xi_1) \in \mathbb{R}_+^2$, then the distribution of $\widetilde{\theta}$ is $\widetilde{F}(\cdot) = F((\cdot - \xi_0)/\xi_1)$. To justify the observational equivalence, we need to show that $(D^{(1)}, D^{(2)}, C, \overline{B}, \underline{B}) = (\widetilde{D}^{(1)}, \widetilde{D}^{(2)}, \widetilde{C}, \widetilde{\overline{B}}, \widetilde{\underline{B}})$, and that the equality (7) holds under the structure $\widetilde{\mathcal{S}}$. Let $\widetilde{\theta}_l = \xi_0 + \xi_1\theta_l$ and $\widetilde{\theta}_u = \xi_0 + \xi_1\theta_u$, then for any $\widetilde{\theta}$, $\widetilde{\theta} \le \widetilde{\theta}_l$ is equivalent to $\theta \le \theta_l$, which implies that $\widetilde{D}^{(1)} = D^{(1)}$. Similarly, we have $\widetilde{D}^{(2)} = D^{(2)}$. Note that

$$\widetilde{\psi}'(\widetilde{e}^*) = \widetilde{H}'(\widetilde{\theta} - \widetilde{e}^*) \Rightarrow \psi'[(\widetilde{e}^* - \xi_0)/\xi_1] = H'[(\widetilde{\theta} - \widetilde{e}^*)/\xi_1],$$

which leads to $\widetilde{e}^*(\widetilde{\theta}) = \xi_0 + \xi_1 e^*(\theta)$. For those with FP contracts,

$$\begin{aligned}
\widetilde{C} &= \widetilde{H}(\widetilde{\theta} - \widetilde{e}^*(\widetilde{\theta}^*)) = H[(\xi_1\theta - \xi_1 e^*(\theta^*))/\xi_1] = H(\theta - e^*(\theta^*)) = C, \\
\widetilde{\overline{B}} &= \widetilde{H}(\widetilde{\theta}_u - \widetilde{e}^*(\widetilde{\theta}_u)) + \widetilde{\psi}(\widetilde{e}^*(\widetilde{\theta}_u)) = H((\xi_1\theta_u - \xi_1 e^*(\theta_u))/\xi_1) + \psi(\xi_1 e^*(\theta^*)/\xi_1) = \overline{B}, \\
\widetilde{\underline{B}} &= r[\widetilde{H}(\widetilde{\theta}_l - \widetilde{e}^*(\widetilde{\theta}_l)) + \widetilde{\psi}(\widetilde{e}^*(\widetilde{\theta}_l))] + (1-r)\overline{B} \\
&= r[H(\theta_l - e^*(\theta_l)) + \psi(e^*(\theta_l))] + (1-r)\overline{B} = \underline{B}.
\end{aligned}$$

For those associated with CR contracts, since $\widetilde{C} = \widetilde{H}(\widetilde{\theta}) = H(\widetilde{\theta}/\xi_1) = H((\xi_0 + \xi_1\theta)/\xi_1)$, then $\widetilde{C} = C$ is equivalent to $\xi_0 = 0$ by noting that $C = H(\theta)$. In what follows, we just need to consider that $\widetilde{\theta} = \xi_1\theta$. Since

$$\widetilde{f}(\widetilde{\theta}_j^*) = \frac{\partial \widetilde{F}(\widetilde{\theta}_j^*)}{\partial \widetilde{\theta}_j^*} = \frac{\partial F(\widetilde{\theta}_j^*/\xi_1)}{\partial \widetilde{\theta}_j^*} = f(\widetilde{\theta}_j^*/\xi_1)/\xi_1 = f(\theta_j^*)/\xi_1, \ \ j = 1, 2$$

we have

$$\left(1 - \frac{\alpha}{1+\lambda}\right) \frac{\widetilde{F}(\widetilde{\theta}_u) - \widetilde{F}(\widetilde{\theta}_l)}{\widetilde{f}(\widetilde{\theta}_u)} = \xi_1 \left(1 - \frac{\alpha}{1+\lambda}\right) \frac{F(\theta_u) - F(\theta_l)}{f(\theta_u)} = \xi_1 \frac{H(\theta_u) - \overline{B}}{H'(\theta_u - e^*(\theta_u, w))}$$

and

$$\frac{\widetilde{H}(\widetilde{\theta}_u) - \widetilde{\overline{B}}}{\widetilde{H}'(\widetilde{\theta}_u - \widetilde{e}^*(\widetilde{\theta}_u))} = \frac{\widetilde{H}(\widetilde{\theta}_u) - \widetilde{\overline{B}}}{\widetilde{\psi}'(\widetilde{e}^*(\widetilde{\theta}))} = \frac{H(\theta_u) - \overline{B}}{\psi'(e^*)/\xi_1} = \xi_1 \frac{H(\theta_u) - \overline{B}}{H'(\theta_u - e^*(\theta_u))}.$$

Hence,

$$\left(1 - \frac{\alpha}{1+\lambda}\right) \frac{\widetilde{F}(\widetilde{\theta}_u) - \widetilde{F}(\widetilde{\theta}_l)}{\widetilde{f}(\widetilde{\theta}_u)} = \frac{\widetilde{H}(\widetilde{\theta}_u) - \widetilde{\overline{B}}}{\widetilde{H}'(\widetilde{\theta}_u - \widetilde{e}^*(\widetilde{\theta}_u))}.$$

This completes the proof.

Proof of Proposition 2. In this proof, we discuss in details the assumptions required to identify the distribution of E^* using Schennach and Hu (2013), as well as other related issues.

In addition to Assumption 2, we need to impose the following restrictions to achieve identification.
(a) The characteristic function of V_1 and V_2 do not vanish anywhere.
(b) The distribution of E^* admits a uniformly bounded density $f_{E^*}(e)$ with respect to the Lebesgue measure that is supported on an interval (which may be infinite).
(c) The function $m(\cdot)$ is continuously differentiable over the interior of the support of E^*.

(d) The set $\chi \equiv \{e : m(e) = 0\}$ has at most a finite number of elements $e_1, \cdots, e_s$ If χ is nonempty, $f_{E^*}(e)$ is continuous and nonvanishing in a neighborhood of each e_l, $l = 1, \cdots s$.

Part (a) is a widely used assumption in the literature of measurement errors. Most of the commonly encountered distributions satisfy this condition, with the notable exceptions being the uniform and the triangular distributions. Parts (b)-(c) are standard smoothness constraints. Part (d) states that we allows for non-monotone function $m(\cdot)$, but rules out functions that are constant over an interval (not reduced to a point) or that exhibit an infinite number of oscillations. Nevertheless, this condition is sufficiently flexible to encompass most specifications of practical interest.

Given Assumption 2 and the conditions (a)-(d), we have the following results on the identification of $f_E(\cdot)$.

1. If $m(\cdot)$ is not of the form $m(e) = a + b\ln(exp(ce) + d)$ for some constants $a, b, c, d \in \mathbb{R}$. Then, $f_E(e)$ and $m(e)$ are nonparametrically identified.

2. If $m(\cdot)$ is linear, i.e., of the form above with $d = 0$, $f_E(e)$ and $m(e)$ are identified.

Note that if $m(\cdot)$ is linear, Schennach and Hu (2013) show that neither $f_{E^*}(\cdot)$ nor $m(\cdot)$ is identified if and only if E^* is normally distributed and either V_1 or V_2 can be decomposed as a summation of two variables with one of them being normally distributed. However, in our setting, the effort e is assumed to be positive and it cannot be normally distributed. Thus both $f_{E^*}(e)$ and $m(e)$ are identified. According Theorem 1 in Schennach and Hu (2013), the only scenario where we cannot identify $f_{E^*}(e)$ and $m(e)$ is that (i) $m(\cdot)$ is of the form $m(e) = a + b\ln(exp(ce) + d)$ with $d \neq 0$, (ii) E^* has a density of the form $f_{E^*}(e) = A\exp(-Bexp(Ce) + CDe)(exp(De) + G)^{-W}$ where $C \in \mathbb{R}$, $A, B, D, G, W \in [0, \infty)$ and (iii) V_2 can be written as a summation of two random variables with one of them being a Type I extreme value variable.

Proof of Proposition 3. Before we present the proof, it is useful to point out several properties of the model under Assumptions 3 and 4. First, the range of realized costs under FP for firms with $W = \varpi_1$ and $W = \varpi_2$, respectively, are as follows:

$$\begin{aligned} &[\underline{C}, C_1] \equiv [C_1(0), C_1(1)], && W = \varpi_1, \\ &[\underline{C}, C_2] \equiv [C_2(0), C_2(1)] = [C_2(0), C_2(p_1/p_2)] \cup (C_2(p_1/p_2), C_2(1)], && W = \varpi_2, \end{aligned}$$

where $C_1(0) = H(\underline{\theta} - e_1(\underline{\theta})) = C_2(0) = H(\underline{\theta} - e_2(\underline{\theta}))$ because Assumption 3 imposes that $e_1(\underline{\theta}) = e_2(\underline{\theta}) = \underline{e}$. Second, for any quantile $\tau \in (0, p_1]$, the optimal effort and realized cost under $W = \varpi_1$ and $W = \varpi_2$ satisfy

$$e_1(\theta(\tau)) < e_2(\theta(\tau)), \quad c_1(\theta(\tau)) > c_2(\theta(\tau))C_1(\tau/p_1) > C_2(\tau/p_2). \tag{A.11}$$

Or, equivalently,

$$E_1(\tau/p_1) < E_2(\tau/p_2), \quad C_1(\tau/p_1) > C_2(\tau/p_2). \tag{A.12}$$

The first inequality can be proved below by contradiction. Note that $\psi_1'(e) > \psi_2'(e)$ for any $e > \underline{e}$ and (A.2) imply

$$H'(\theta(\tau) - e_1(\theta(\tau))) = \psi_1'(e_1(\theta(\tau))) > \psi_2'(e_1(\theta(\tau))).$$

Suppose $e_1(\theta(\tau)) \geq e_2(\theta(\tau))$, then $H''(\cdot) > 0$ further implies

$$\begin{aligned} \psi_1'(e_1(\theta(\tau))) &= H'(\theta(\tau) - e_1(\theta(\tau))) \\ &\leq H'(\theta(\tau) - e_2(\theta(\tau))) = \psi_2'(e_2(\theta(\tau))) \leq \psi_1'(e_2(\theta(\tau))). \end{aligned} \tag{A.13}$$

Under the assumption $\psi''(\cdot) > 0$, the equation above implies that $e_1(\theta(\tau)) < e_2(\theta(\tau))$, which contradicts $e_1(\theta(\tau)) \geq e_2(\theta(\tau))$. Therefore, $e_1(\theta(\tau)) < e_2(\theta(\tau))$. As a result, $C_1(\tau/p_1) > C_2(\tau/p_2)$ because the realized cost $C_j = H(\theta - e_j(\theta))$. From the second inequality, we obtain that for any $\eta < p_2, C_1(\eta) > C_2(\eta.p_1/p_2)$.

Next, we present the details of the proof. We first choose any realized cost $t \in [C_1(0), C_1(1)]$, then there will be a unique quantile $\tau_0 \in [0, p_1]$ of cost C corresponds to t

$$\tau_0(t) = \begin{cases} 0, & t = C_1(0) = \underline{C}; \\ C_1^{-1}(t), & t \in (C_1(0), C_1(1)]. \end{cases} \tag{A.14}$$

Applying (A.11), we have

$$t = C_1(\tau_0(t)) > C_2(\tau_0(t) \cdot p_1/p_2).$$

Similarly, there exists a unique $\tau_1 \in (0, \tau_0(t))$ such that

$$C_1(\tau_1(t)) = C_2(\tau_0(t).p_1/p_2) > C_2(\tau_1(t).p_1/p_2).$$

We repeat this procedure to obtain the following inequlities:

$$\begin{aligned} t = C_1(\tau_0(t)) &> C_2(\tau_0(t) \cdot p_1/p_2) = C_1(\tau_1) > C_2(\tau_1 \cdot p_1/p_2) \\ &... \\ = C_1(\tau_k(t)) &> C_2(\tau_k(t) \cdot p_1/p_2) = C_1(\tau_{k+1}(t)) \\ &... \end{aligned}$$

So it ends up with a bounded and decreasing sequence $\{\tau_k(t)\} \subseteq (0, 1)$, for which we know a unique limit exists, denoted by $\underline{\tau}(t)$. As $\tau_k(t) \to \underline{\tau}(t)$, we have $C_1(\tau_k(t)) \to$

$C_1(\underline{\tau}(t))$, and $C_2(\tau_k(t) \cdot p_1/p_2) \to C_2(\underline{\tau}(t) \cdot p_1/p_2)$. Note that as $\tau_k(t)$ so picked, $C_2(\tau_k(t) \cdot p_1/p_2) = C_1(\tau_{k+1}(t))$. Taking limit on both sides implies

$$C_2(\underline{\tau}(t) \cdot p_1/p_2) = C_1(\underline{\tau}(t)),$$

which can be only true for $\underline{\tau}(t) = 0$. Therefore, it has to be the case that $\tau_k(t) \to 0$.

Now it is readily to derive the iteration equation

$$\begin{aligned} H^{-1}(t) &= H^{-1}(C_1(\tau_0(t))) \\ &= H^{-1}(C_2(\tau_0(t) \cdot p_1/p_2)) + \Delta\tilde{E}(\tau_0(t) \cdot p_1) \\ &= H^{-1}(C_1(\tau_1(t)) + \Delta\tilde{E}(\tau_0(t) \cdot p_1) \\ &= H^{-1}(C_2(\tau_1(t) \cdot p_1/p_2)) + \Delta\tilde{E}(\tau_1(t) \cdot p_1) + \Delta\tilde{E}(\tau_0(t) \cdot p_1) \\ &= \cdots \\ &= H^{-1}(C_2(\tau_m(t) \cdot p_1/p_2)) + \sum\nolimits_{k=0}^{m} \Delta\tilde{E}(\tau_k(t) \cdot p_1). \end{aligned} \tag{A.15}$$

Rearranging terms yields

$$\sum\nolimits_{k=0}^{m} \Delta\tilde{E}(\tau_k(t) \cdot p_1) = H^{-1}(x) - H^{-1}(C_2(\tau_m(t) \cdot p_1/p_2)).$$

Taking limit on both side of the equation above yields

$$\sum\nolimits_{k=0}^{\infty} \Delta\tilde{E}(\tau_k(t) \cdot p_1) = H^{-1}(x) - H^{-1}(C_2(0)).$$

Or, equivalently,

$$\sum\nolimits_{k=0}^{\infty} \Delta\tilde{E}(\tau_k(t) \cdot p_1) = H^{-1}(x) - (\underline{\theta} - \underline{e}).$$

Rearranging terms again yields

$$H^{-1}(t) = \underline{\theta} - \underline{e} + \sum\nolimits_{k=0}^{\infty} \Delta\tilde{E}(\tau_k(t) \cdot p_1).$$

Therefore, $H^{-1}(\cdot)$ is identified for $C \in [C_1(0), C_1(1)]$ because both the sequence $\{\tau_k(t)\}$ and each $\Delta\tilde{E}(\tau_k(t) \cdot p_1) = E_2(\tau_k(t) \cdot p_1/p_2) - E_1(\tau_k(t))$ are identified. This completes the proof.

Proof of Corollary 1. The proof of Corollary 1 is similar to that of Proposition 3. The main difference of identification in Corollary 1 is to first identify the intersection point e_c which corresponds to the intersection point of the two cost distributions under $W = \varpi_1$ and $W = \varpi_2$. Under Assumption 4, it is easy to show that there exists a $\theta_c \in [\underline{\theta}, \theta_u^1]$ such that $E_1(\tau_c/p_1) = E_2(\tau_c/p_2) = e_c$, where τ_c satisfies $\theta_c = \theta(\tau_c)$ due to the one-to-one mapping between cost and type. And, τ_c is identified by $C_1(\tau_c/p_1) =$

$C_2(\tau_c/p_2)$ since the two distribution functions of C_1 and C_2 intersect only once at the cost quantile corresponding to $\theta(\tau_c)$. As a result, e_c is also identified. Obviously, $H^{-1}(t)$ can be identified for any t on its support as long as $\theta(\tau_c)$ can be recovered due to the following identification equation similar to (16)

$$H^{-1}(t) = \theta(\tau_c) - e_c + \sum_{k=0}^{\infty} \Delta\tilde{E}(\tau_k(t).p_1).$$

Following the identification argument in Proposition 3, we obtain

$$H^{-1}(C_1(0)) = \theta(\tau_c) - e_c + \sum_{k=0}^{\infty} \Delta\tilde{E}(\tau_k(C_1(0)).p_1).$$

On the other hand,

$$H^{-1}(C_1(0)) = \underline{\theta} - e_1(0).$$

Therefore, $\theta(\tau_c)$ is identified as

$$\theta(\tau_c) = \underline{\theta} - e_1(0) + e_c - \sum_{k=0}^{\infty} \Delta\tilde{E}(\tau_k(C_1(0)).p_1).$$

Proof of Proposition 4. The proposed estimators $\{\hat{\theta}_i\}_{i=1}^n$, $\widehat{F}_\theta(\cdot)$, $\widehat{\alpha/(1+\lambda)}$, $\widehat{\psi}(\cdot)$, and $\hat{\delta}$ can all be viewed as plug-in estimators based on the estimated cost function $\widehat{H}(\cdot)$. So the consistency of $\widehat{H}(\cdot)$ is the key to guarantee consistency of all other estimators. Under the parameterization $H(\cdot;\beta)$ with it being continuous in β, $H(t;\beta)$ is consistently estimated by $H(t;\hat{\beta})$ for any given $t \in [\underline{\theta} - \underline{e}, \overline{\theta}]$ if $\hat{\beta}$ is consistent for β, according to the continuous mapping theorem (CMT). In what follows in the proof of Proposition 4, we focus on showing the consistency of $\hat{\beta}$.

Define

$$Q(\beta) = \sum_{l=1}^{L} \left[H^{-1}(C_2(a_l \cdot \frac{p_1}{p_2}); \beta) - H^{-1}(C_1(a_l); \beta) + E_2(a_l \cdot \frac{p_1}{p_2}) - E_1(a_l) \right]^2. \quad \text{(A.16)}$$

and

$$Q_n(\beta) = \sum_{l=1}^{L} \left[H^{-1}(\widehat{C}_2(a_l \cdot \frac{\hat{p}_1}{\hat{p}_2}); \beta) - H^{-1}(\widehat{C}_1(a_l); \beta) + \widehat{E}_2(a_l \cdot \frac{\hat{p}_1}{\hat{p}_2}) - \widehat{E}_1(a_l) \right]^2. \quad \text{(A.17)}$$

Then the true value of β, denoted by β_0, is equivalently characterized as

$$\beta_0 = \underset{\beta: H(\underline{\theta} - \underline{e}; \beta) = \underline{c}}{\operatorname{argmin}} \ Q(\beta). \quad \text{(A.18)}$$

And the estimator is

$$\hat{\beta} = \underset{\beta: H(\underline{\theta}-\underline{e};\beta)=\underline{c}}{\operatorname{argmin}} Q_n(\beta). \tag{A.19}$$

As shown above, $\hat{\beta}$ is an extremum type estimator.

When β is identified (which is a condition required in Proposition 4),

$$\inf_{\beta \in \mathcal{B}: d(\beta,\beta_0) \geq \epsilon} Q(\beta) > 0 = Q(\beta_0) \tag{A.20}$$

for every $\epsilon > 0$, due to the compactness of $\{\beta \in \mathcal{B} : d(\beta, \beta_0) \geq \epsilon\}$ and continuity of $Q(\cdot)$, the latter of which follows from the boundedness and continuity of $H(\cdot)$.

Also, the compactness of $\mathcal{B}$, the continuity of $H(\cdot)$, and the consistency of $\{\hat{p}_1, \hat{p}_2, \widehat{C}_1(a_l), \widehat{C}_2(a_l \cdot \frac{p_1}{p_2}), \widehat{E}_1(a_l), \widehat{E}_2(a_l \cdot \frac{p_1}{p_2})\}_{l=1}^{L}$ guarantee that

$$\sup_{\beta \in \mathcal{B}} |Q_n(\beta) - Q(\beta)| \xrightarrow{p} 0. \tag{A.21}$$

According to 5.7 Theorem in van der Vaart (1999), (A.20) and (A.21) imply that $\hat{\beta} \xrightarrow{p} \beta_0$, which completes the proof.

Identification and estimation under constant optimal effort: Denote by $\overline{c}_F$ the upper bound of costs associated with CF choices, then

$$\overline{c}_F = H(\theta_u - e^*) = \beta(\theta_u - e^*). \tag{A.22}$$

Using the optimal condition $\overline{b} = H(\theta_u - e^*) + \psi(e^*)$ in (7), one obtains

$$\psi(e^*) = \overline{b} - \overline{c}_F$$

and hence we identify the optimal effort level as

$$e^* = (\overline{b} - \overline{c}_F)^{1/2}.$$

Denote by $\underline{c}_R$ the lower bound of costs associated with CC choices, then this realized cost

$$\underline{c}_R = H(\theta_u) = \beta\theta_u. \tag{A.23}$$

Combining (A.22) with (A.23), we can identify the parameter β as

$$\beta = (\underline{c}_R - \overline{c}_F)/(\overline{b} - \overline{c}_F)^{1/2}.$$

Once β is identified, we can recover the innate cost as follows.

$$\theta = \begin{cases} c(\bar{b} - \bar{c}_F)^{1/2}/(\underline{c}_R - \bar{c}_F) + (\bar{b} - \bar{c}_F)^{1/2}, \text{ for FP contracts,} \\ \\ c(\bar{b} - \bar{c}_F)^{1/2}/(\underline{c}_R - \bar{c}_F), \text{ for CR contracts.} \end{cases} \tag{A.24}$$

The distribution of innate costs $F(\cdot)$ is then identified based on the recovered θ. The discount factor δ and the bargaining power α can be identified using (7).

The constructive identification strategy above suggests a semiparametric estimation procedure to estimate all the model primitives. We omit the details in the paper.