A Structural Analysis of Simple Contracts*

Yonghong An† Shengjie Hong‡ Daiqiang Zhang§

Abstract

This paper provides an econometric framework for analyzing simple contracts where an agent chooses between a fixed-price option and a cost-reimbursement option provided by a principal in each contracting period during possibly multiple periods. First, we propose a consistent procedure for testing the null hypothesis of a corresponding cost function being linear, which is widely assumed for tractability in the literature. Motivated by the rejection of such a null based on our empirical data, next we establish nonparametric identification, without restricting the cost function to be linear, for all model primitives conditioned on the agent exerting nonzero effort. These primitives include agent's cost and disutility functions, distribution of agent efficiency type, and parameters that characterize agent's bargaining power and intertemporal preference. Moreover we propose a consistent procedure to implement the identification results for estimation. In our empirical study, we find strong evidence against linearity of the cost function. The importance of this empirical finding is further evidenced by a welfare analysis, which shows the welfare assessment to be sensitive to the specification of cost function.

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). This paper supersedes a previous working paper "Econometrics of Multi-Period Simple Contracts". All remaining errors are ours.

†Department of Economics, Texas A&M University, College Station, TX 77843; email: y.an@tamu.edu.

‡School of Economics and Management, 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 understanding informational asymmetries and incentives, contract theory has been a rapidly developing field in economics for the past three decades. Early studies had been focused on *complex optimal contracts*, in the spirit of Laffont and Tirole (1986), where the payment scheme to an agent is specified as a function of both the agent's observed cost and his unobserved type. More recently, *simple contracts*, which are also known as (a.k.a.) *simple menu contracts* in the literature, have attracted a lot of attention. This branch of contracts specify the payment scheme simply as a function of only the agent's observed cost or even as a constant. Not surprisingly, simple contracts are more widely adopted in practice than complex optimal contracts, because their payment schemes are much easier to implement. The theoretic importance of simple contracts has been much recognized. In particular, evidences from both theory and empirical studies have shown that simple contracts can capture a substantial proportion of the surplus that complex optimal contracts would achieve (Rogerson, 2003; D'Haultfœuille and Février, 2020). However, econometric analyses of simple contracts are still very limited.

In this paper, we study simple contracts with two basic options: fixed-price (FP) and cost-reimbursement (CR). The FP option specifies the payment to be a fixed amount regardless of the agent's realized cost. In contrast, the CR option specifies the payment to equal exactly the agent's realized cost. FP-CR menus are widely used in practice, e.g., they are commonly adopted by the U.S. Department of Defense (Rogerson, 1992). They are also used by local authorities in France to contract with operators to provide transportation service(Aaker and Myers, 1987; Gagnepain et al., 2013). Additional 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), among many others. Our goal is to develop a flexible framework for econometric analysis of FP-CR contracts.

First, we construct an economic model to serve as the basis for our structural econometric analyses. We take Gagnepain et al. (2013)'s economic model as our starting point and depart from there by relaxing one of their key assumptions, namely the *linear cost condition* (LCC hereafter). The LCC restricts the agent's cost of fulfilling a contract to be linear in his efficiency type (a.k.a. innate cost). Often assumed for tractability, the LCC also helps to achieve econometric identification when coupled with other simplifying assumption(s), e.g., a normal distribution for the efficiency type as in Gagnepain et al. (2013). However, its empirical validity is questionable. For instance, some empirical evidences suggest that the optimal effort to be monotonic in the efficiency type (see, e.g., Gagnepain and Ivaldi, 2002 and Abito, 2020), as opposed to being constant (i.e., being invariant in the efficiency type) which necessarily holds under the LCC. After establishing the economic model, we specify the econometric setting and focus on two major tasks: to develop a credible inference procedure to assess whether the LCC is consistent with the data, and to develop identification strategies for the model primitives without imposing the LCC.

To assess whether the LCC is consistent with the data, we propose a consistent procedure for testing a null hypothesis that is directly implied by the LCC. The usefulness of the proposed testing procedure is two-fold: (i) The modeling simplicity under the LCC is appealing and appreciated, especially if there is no significant empirical evidence against it. So if the proposed test fails to reject the null, it is acceptable to impose the LCC; (ii) If the null is rejected, it is better to adopt a more credible way of econometric analyses, which does not rely on the LCC.

The other major task is to develop identification strategies for the model primitives without imposing the LCC. These primitives include agent's cost function, agent's disutility function (from exerting cost-reducing effort), distribution of agent's efficiency type, and parameters relating to agent's bargaining power and intertemporal preference. Their identifications are achieved through the following steps: First, by adopting a recent method on measurement error by Schennach and Hu (2013), we recover the distribution of the unobserved optimal effort from the joint distribution of two observable effort-related proxies; Next, we work on identification of the cost function, which is the key to identify all other model primitives. We require the existence of an exclusion variable that is independent from agent's type but affects the disutility from exerting cost-reducing efforts. Our strategy to identify the cost function is to match quantiles of the cost and effort distributions conditioned on different values of the exclusion variable, according to the corresponding quantiles of the type distribution, which are invariant to the variable; Third, with the cost function identified, we exploit the structural link between the agent's efficiency type and the corresponding optimal effort level to recover the type value associated with any given observed cost under FP contracts; Last, based on what have been identified, we identify the remaining primitives. Following the identification strategies, we propose consistent methods to estimate all model primitives. Throughout our analysis, we prioritize on identification. To the best of our knowledge, ours is the first set of positive results on identification of multi-period simple contracts in the literature.

As an empirical illustration, we apply our methods to study transportation procurement contracts in France. The main objectives are to test the widely assumed LCC, and to evaluate how this specification assumption affects welfare assessment. Based on our empirical data, a direct implication from the LCC is reject at a significance level of 1%. This testing result suggests that the LCC fails to adequately describe the cost structure in the French transportation industry. To evaluate how the LCC affects the welfare assessment, we estimate the welfare with and without imposing the LCC then conduct a comparison. We find substantial difference between the two assessments, which suggests the welfare assessment to be sensitive to the specification of cost function. Thus, one needs to be cautious on deciding whether or not to impose linearity, since mis-specification could lead to substantial bias.

Our study contributes to the growing literature on identification of contract models; see, Perrigne and Vuong (2011) and D'Haultfœuille and Février (2020), among others. Perrigne and

Vuong (2011) establish nonparametric identification of a static *complex* contract model tailored from the seminal paper Laffont and Tirole (1986). D'Haultfœuille and Février (2020) show nonparametric *partial identification* of simple contracts using exogenous variations of contracts. We note that the identification strategy of Perrigne and Vuong (2011) is not applicable to our setting, mainly because the one-to-one mapping between observed product price and agent type, a key to achieve identification in Perrigne and Vuong (2011), is unavailable from the simple contracts that we consider. Compared with D'Haultfœuille and Février (2020), our paper differs in both model setup and identification strategies.[1] Our paper is closely related to Gagnepain et al. (2013) in the sense that our economic model of FP-CR contracts adopts a similar principal-agent setup. Besides, we use the same dataset of French public transportation service contracts as they do for our empirical study. Nevertheless, Gagnepain et al. (2013) focus on evaluating the social cost of allowing for renegotiation, where the whole analysis is under the assumption that the LCC holds. While we mainly focus on testing the validity of the LCC and developing identification results without imposing it.

In a broader view, our identification result is related to the econometric analysis on a richer class of principal-agent models featuring moral hazard and adverse selection, where an agent with unknown type is offered a menu of only a few simple contractual options by the principal, such as the insurance models studied by Aryal et al. (2010) and the nonlinear pricing models studied by Luo et al. (2018). In addition, a growing number of papers utilize developments in measurement errors to identify structural models. See Hu (2017) for a recent survey. Our paper also contributes to this category of literature: As far as we know, our study is the first to employ the results in measurement errors to identify contract models.

The rest of the paper is organized as follows. In Section 2 we present the economic model of FP-CR contracts. In Section 3 we specify our econometric setting and propose a formal test for the LCC. In Section 4 we establish the main identification results. In Section 5 we propose consistent estimation procedures. In Section 6 we present an empirical study on French transportation procurement contracts. We conclude the paper in Section 7. Proofs of theorems, lemmas and corollaries are collected in the Appendix.

2 The model

A local authority (the 'principal') wants to procure a public service or project from a firm (the 'agent'). A procurement contract can last for a single period or for multiple periods. A single-period contract sets a payment option, with all associated terms fully specified, from a menu of options offered by the principal. A multi-period contract sets a sequence of fully

[1] In D'Haultfœuille and Février (2020), unlike standard contract models, agent's disutility function of effort is not explicitly modeled. In contrast, we model the disutility function explicitly. Moreover, while D'Haultfœuille and Février (2020) provide very informative partial identification results, we aim at developing regularity conditions for point identification.

specified options, one for each period. Let $T < \infty$ be the total number of contracting periods. A multi-period contract can be signed with *full commitment* or with *renegotiation* permitted. In the latter case, there is a window in between periods t and $t+1$ (i.e., between the end of period t and the beginning of period $t+1$), for each $t = 1, ..., T-1$, during which the principal and agent are allowed to renegotiate to alter options and/or their detailed terms for subsequent periods (i.e., from $t+1$ to T). In contrast, a contract with full commitment prohibits any renegotiation once signed. In our study, we focus on single-period contracts, and two-period ones that permit renegotiation.

2.1 Basic setup

The agent's per period cost c_t is determined by his efficiency type θ (a.k.a. innate cost) and his cost reducing effort e_t, according to a general cost function $H(\cdot)$ as

$$c_t = H(\theta - e_t), \quad \text{for } t = 1, ..., T.$$

Often interpreted in the literature as a measure of management and production skills, θ is randomly drawn from a cumulative distribution function (CDF) $F(\cdot)$, with density $f(\cdot)$, on a bounded support $\Theta = [\underline{\theta}, \bar{\theta}] \subset \mathbb{R}$. θ is the agent's private information and is specified as time-invariant, while $F(\cdot)$ is common knowledge. $e_t \geq 0$ measures the agent's cost-reducing effort in period t, which induces some disutility according to a disutility function $\psi(e_t)$. We impose the following conditions, which are standard in related literature (e.g., Laffont and Tirole, 1993):

Assumption 2.1 *(i) $H(\cdot) \geq 0$, $H'(\cdot) > 0$ and $H''(\cdot) > 0$; (ii) $\psi(\cdot) \geq 0$, $\psi'(\cdot) > 0$, $\psi''(\cdot) > 0$ and $\psi(0) = 0$.*

The agent's payoff in period t is

$$q_t - c_t - \psi(e_t), \text{ with } q_t = \begin{cases} c_t \text{ under CR;} \\ b_t \text{ under FP,} \end{cases}$$

where q_t is the payment he receives from the principal, and b_t is the fixed price associated with the FP option. Given b_t, the agent chooses between FP and CR, and decides an optimal level of effort, to maximize his payoff. Under FP, the optimal level of effort depends on θ in general, hence denoted by $e(\theta)$. We have

$$e(\theta) = \underset{e}{\operatorname{argmin}} \{H(\theta - e) + \psi(e)\}, \tag{2.1}$$

which satisfies the first-order condition (f.o.c.)

$$H'(\theta - e(\theta)) = \psi'(e(\theta)). \tag{2.2}$$

Consequently, under FP with a fixed price b_t, the maximum payoff achievable is

$$u_F(\theta, b_t) \equiv b_t - H(\theta - e(\theta)) - \psi(e(\theta)). \tag{2.3}$$

Since the CR option provides no incentive for any cost reducing effort, the optimal effort level under CR would be zero regardless of θ. As a result, under the normalizing condition $\psi(0) = 0$ in Assumption 2.1, the payoff under CR would always be zero. So the agent will choose FP (with b_t) over CR in period t if and only if $u_F(\theta, b_t) \geq 0$. With optimal choices of option and effort level, his per period payoff would be

$$u(\theta, b_t) \equiv \max\{u_F(\theta, b_t), 0\}. \tag{2.4}$$

The principal is primarily concerned with social welfare from fulfillment of the contract. In generic terms, the social welfare is defined as

$$SW = S - (1+\lambda) Q + \alpha U \tag{2.5}$$

where S is the gross surplus generated by the procured service, Q is the principal's payment to the agent, and U is the agent's payoff. The way the principal raises the payment funds Q, usually by imposing distortion tax, typically leads to a dead-weight loss. Such a dead-weight loss is captured by λ. α measures the agent's bargaining power. From an alternative view, α can be interpret as reflecting the local authority's political preferences. The definition (2.5) above dates back to Baron and Myerson (1982), and have been widely used in the contract literature.[2]

Suppose, in period t, the principal makes a payment q_t to the agent, and the agent's payoff ends up being u_t. According to the general definition (2.5), the per period social welfare is

$$\pi_t = S - (1+\lambda) q_t + \alpha u_t,$$

Anticipating the agent's choices of option and effort level in response to a fixed price offering of b_t, the principal sets b_t to maximize expected social welfare, which we explain in details as we discuss equilibrium properties in Sections 2.2 and 2.3. Before we proceed, we note that the following per period patterns hold in both the single-period and two-period settings:

Lemma 2.1 *Let Assumption 2.1 hold. (i) Under CR, the agent exerts zero effort regardless of θ, and the cost is realized as $c_t = H(\theta)$; (ii) Under FP, the optimal effort $e(\theta)$ is strictly increasing in θ, with $e'(\theta) \in (0,1)$; (iii) Moreover, under FP, both $\theta - e(\theta)$ and $H(\theta - e(\theta))$ (i.e., the realized cost c_t) are strictly increasing in θ.*

2.2 Equilibrium for single-period contracts

In the single-period setting, we drop the subscript t for indicating a specific period. In equilibrium, for any fixed price b she may offer, the principal correctly anticipates the agent's choices of option and effort level in response conditional on θ, as specified by (2.1), as well as

[2]For more discussions on λ and α, see Baron and Myerson, 1982, Baron, 1988 and Gagnepain et al., 2013. Also, S is assumed to be sufficiently large to guarantee the desirability of the procured service or project.

his resulting payoff, as specified by (2.4). She predicts the social welfare resulting from offering b, conditional on θ, to be

$$\tilde{\pi}(\theta, b) = S - (1+\lambda)\left\{\mathbf{1}\left(u_{_F}(\theta, b) \geq 0\right) \cdot b + \left[1 - \mathbf{1}\left(u_{_F}(\theta, b) \geq 0\right)\right] \cdot H(\theta)\right\} + \alpha u(\theta, b). \quad (2.6)$$

So she sets an optimal fixed price b^* to maximize the *expected social welfare* defined as

$$\pi(b) \equiv \mathbb{E}_\theta\left[\tilde{\pi}(\theta, b)\right] = \int_\Theta \tilde{\pi}(\theta, b)\, dF(\theta). \quad (2.7)$$

Proposition 2.1 (Single-period equilibrium) *Consider a single-period contract. Let Assumption 2.1 hold. In equilibrium, the following holds:*
(i) There is a unique cut-off value $\theta^* \in \left[\underline{\theta}, \bar{\theta}\right]$ *such that the agent chooses FP if* $\theta \leq \theta^*$*, i.e., being more efficient, and chooses CR otherwise.*
(ii) The principal offers the optimal fixed price $b^* = \underset{b}{\operatorname{argmax}}\, \pi(b)$*.* b^* *and* θ^* *satisfy*

$$b^* = H\left(\theta^* - e\left(\theta^*\right)\right) + \psi\left(e\left(\theta^*\right)\right), \quad (2.8)$$

$$\left(1 - \frac{\alpha}{1+\lambda}\right)\frac{F\left(\theta^*\right)}{f\left(\theta^*\right)} = \frac{H\left(\theta^*\right) - b^*}{H'\left(\theta^* - e\left(\theta^*\right)\right)}. \quad (2.9)$$

2.3 Equilibrium for two-period contracts

For a two-period FP-CR contract, the principal can offer four possible combinations of per period options. The option-menu is as follows.

$$\begin{cases} \textbf{FF}\text{: FP for both periods, with payments } \left(b_1^{FF}, b_2^{FF}\right); \\ \textbf{CF}\text{: CR for period-1 and FP for period-2, with payments } \left(c_1, b_2^{CF}\right); \\ \textbf{CC}\text{: CR for both periods, with payments } \left(c_1, c_2\right); \\ \textbf{FC}\text{: FP for period-1 and CR for period-2, with payments } \left(b_1^{FC}, c_2\right). \end{cases}$$

We ignore the fourth option, FC, as it would never appear in equilibrium.[3]

Use generic notations u_t and π_t for the payoff and social welfare in period t, respectively, for $t = 1, 2$. The *intertemporal payoff* and *intertemporal social welfare* are defined respectively as

$$\gamma u_1 + (1-\gamma)\, u_2 \text{ and } \gamma \pi_1 + (1-\gamma)\, \pi_2,$$

where $\gamma \equiv 1/(1+\delta)$ (with δ being the discount factor) is a measure of the relative importance of the first period, following Laffont and Tirole (1990)'s interpretation.

[3]This is because, when offered a vector of fixed prices $\left(b_1^{FF}, b_2^{FF}, b_2^{CF}, b_1^{FC}\right)$ that maximizes the principal's objective function (to be specified soon for the two-period setting), the agent would always prefer one of the first three options (i.e., FF, CF and CC) to FC regardless of θ.

Given a vector of fixed prices $\mathbf{b} = \left(b_1^{FF}, b_2^{FF}, b_2^{CF}\right)'$, the maximum agent's payoffs under FF, CF and CC are

$$u_{FF}(\theta, \mathbf{b}) \equiv \gamma u_F\left(\theta, b_1^{FF}\right) + (1-\gamma)\, u_F\left(\theta, b_2^{FF}\right),$$
$$u_{CF}(\theta, \mathbf{b}) \equiv (1-\gamma)\, u_F\left(\theta, b_2^{CF}\right) \text{ and } u_{CC}(\theta, \mathbf{b}) \equiv 0,$$

respectively, where $u_F(\theta, \cdot)$ is defined in (2.3). These maximums are achieved by making an effort of level $e(\theta)$ (as specified in (2.1)) for any FP period, and zero effort for any CR period. Based on θ and $\mathbf{b}$, the agent would choose option $J \in \{FF, CF, CC\}$ that yields the largest $u_J(\theta, \mathbf{b})$ to maximize his intertemporal payoffs. So the agent's choice of option, denoted by $J(\theta, \mathbf{b})$, is characterized as

$$J(\theta, \mathbf{b}) = \underset{J \in \{FF, CF, CC\}}{\operatorname{argmax}} \; u_J(\theta, \mathbf{b}).$$

In equilibrium, correctly anticipating the agent's choices of option and effort level as described above, the principal predicts the intertemporal social welfare conditional on θ to be

$$\tilde{\pi}_{int}(\theta, \mathbf{b}) = S - (1+\lambda)\, q_{J(\theta,\mathbf{b})}(\theta, \mathbf{b}) + \alpha u_{J(\theta,\mathbf{b})}(\theta, \mathbf{b}) \tag{2.10}$$

for any given vector of fixed prices $\mathbf{b} = \left(b_1^{FF}, b_2^{FF}, b_2^{CF}\right)'$, where

$$q_J(\theta, \mathbf{b}) \equiv \begin{cases} \gamma b_1^{FF} + (1-\gamma)\, b_2^{FF}, \text{ under FF;} \\ \gamma H(\theta) + (1-\gamma)\, b_2^{CF}, \text{ under CF;} \\ H(\theta), \text{ under CC} \end{cases} \tag{2.11}$$

is the intertemporal payment she needs to raise under option J, for $J \in \{FF, CF, CC\}$. In equilibrium, the principal sets an optimal vector of fixed prices $\mathbf{b}^*$ to maximize the *expected intertemporal social welfare* defined as

$$\pi_{int}(\mathbf{b}) \equiv \mathbb{E}_\theta\left[\tilde{\pi}_{int}(\theta, \mathbf{b})\right] = \int_\Theta \tilde{\pi}_{int}(\theta, \mathbf{b})\, dF(\theta), \tag{2.12}$$

i.e., $\mathbf{b}^* = \underset{\mathbf{b}}{\operatorname{argmax}}\, \pi_{int}(\mathbf{b})$.

As mentioned earlier, we distinguish between two types of two-period contracts: (i) Contracts with *full commitment*; (ii) Contracts permitting *renegotiation*. Our focus is on the latter, which also fits our empirical study. Nevertheless, we first discuss the former briefly, for a better understanding of the latter.

Full commitment

A contract with full commitment prohibits any adjustment of initial arrangement during implementation. As pointed out by Laffont and Tirole (1990), the two-period optimal contract with full commitment is simply a twice-repeated version of the single-period one. Hence, it holds, for any optimal vector of fixed prices $\mathbf{b}^* = \left(b_1^{FF*}, b_2^{FF*}, b_2^{CF*}\right)' \in \underset{\mathbf{b}}{\operatorname{argmax}}\, \pi_{int}(\mathbf{b})$, that

$$b_2^{CF*} \leq b_1^{FF*} = b_2^{FF*} = b^* = \underset{b}{\operatorname{argmax}}\, \pi(b). \tag{2.13}$$

There is a unique cut-off value $\theta^* \in \left[\underline{\theta}, \bar{\theta}\right]$ such that the agent chooses FF if $\theta \leq \theta^*$, and CC otherwise.[4] For b^* and θ^*, (2.8) and (2.9) in Proposition 2.1 still hold.

Renegotiation

Being observable to the principal, the agent's choice of option and realized cost in the first contracting period reveal partial information on θ. The principal may utilize this information to update her belief on the distribution of θ, which may trigger potential renegotiation in between periods in executing a multi-period contract. Here, we adopt the *limited updating* specification as in Gagnepain et al. (2013): The principal's updating belief is based on information revealed by only the agent's choice of option, but not the realized cost. The principal's initial belief is $F(\cdot)$ with support $\left[\underline{\theta}, \bar{\theta}\right]$. For a given fixed price b, let $\theta(b)$ be the corresponding cut-off type such that an agent of type $\theta(b)$ would be indifferent between FP and CR, i.e.,

$$b = H(\theta(b) - e(\theta(b))) + \psi(e(\theta(b))).$$

(It follows immediately from (2.8) that $\theta^* = \theta(b^*)$.) The agent's choosing FP (with b) over CR reveals that $\theta \leq \theta(b)$, and consequently leads to an updated belief of the principal as follows

$$\underline{F}_b(\cdot) \equiv F(\cdot)/F(\theta(b)), \text{ with support } [\underline{\theta}, \theta(b)],$$

which is $F(\cdot)$ truncated on $[\underline{\theta}, \theta(b)]$. Based on $\underline{F}_b(\cdot)$, the principal reformulates her expected per period social welfare as $\underline{\pi}_b(\cdot) \equiv \int_\Theta \tilde{\pi}(\theta, \cdot)\, \underline{F}_b(\theta)$. Similarly, choosing CR over FP (with b) leads to an updated belief of

$$\bar{F}_b(\cdot) \equiv \frac{F(\cdot) - F(\theta(b))}{1 - F(\theta(b))}, \text{ with support } \left[\theta(b), \bar{\theta}\right],$$

which is $F(\cdot)$ being truncated on $\left[\theta(b), \bar{\theta}\right]$. Based on $\bar{F}_b(\cdot)$, the principal reformulates her expected per period social welfare as $\bar{\pi}_b(\cdot) \equiv \int_\Theta \tilde{\pi}(\theta, \cdot)\, \bar{F}_b(\theta)$. As long as $\theta(b) = \left(\underline{\theta}, \bar{\theta}\right)$, it can be shown that $\underline{\pi}_b(\cdot)$ is maximized at a value smaller than b^*, and that $\bar{\pi}_b(\cdot)$ is maximized at a value larger than b^*,

To further illustrate how renegotiation may happen, suppose the principal and agent sign an initial contract with fixed prices $\mathbf{b}^*$ as in (2.13), which would be the equilibrium profile with full commitment. Consider the following cases: (i) Suppose the agent chooses FF. Realizing that the agent is of more efficiency type (whose θ is distributed according to $\underline{F}_{b^*}(\cdot)$), the principal

[4] Note that, with full commitment, CF would not be selected in equilibrium.

would want to renegotiate for a lower second period fixed price $b_2^{FF} < b_2^{FF*} = b^*$. However, the agent would not agree on a lower fixed price, so renegotiation would not actually happen in this case; (ii) Now suppose the agent chooses CC. Realizing that the agent is of less efficiency type (distributed according to $\bar{F}_{b^*}(\cdot)$), the principal would want to renegotiate for a higher second period fixed price $b_2^{CF} > b^* \geq b_2^{CF*}$ to provide more incentive for the agent to select FP and revamp efficiency in period-2. Since a higher fixed price is always welcomed by the agent, a renegotiated second-period fixed price can be possibly agreed on in this case.

According to the *renegotiation-proof principle*, any agreement of $\{b_1^{FF}, b_2^{FF}, b_2^{CF}\}$, of which the period-2 continuation $\{b_2^{FF}, b_2^{CF}\}$ is superseded by the renegotiated $\{\tilde{b}_2^{FF}, \tilde{b}_2^{CF}\}$, could be replaced by the agreement of $\{b_1^{FF}, \tilde{b}_2^{FF}, \tilde{b}_2^{CF}\}$ under which no renegotiation would actually happen. For this reason, we focus on *renegotiation-proof profiles*, which is a standard arrangement in the contract literature and causes no loss of generality. Proposition 2.2 below characterizes the equilibrium.

Proposition 2.2 (Two-period equilibrium) *Consider a two-period contract that permits renegotiation. Let Assumption 2.1 hold. In equilibrium, the following holds:*
(i) There are two cut-off values $\theta_1^*, \theta_2^* \in [\underline{\theta}, \bar{\theta}]$ with $\theta_1^* < \theta_2^*$ such that the agent chooses FF if $\theta \leq \theta_1^*$, chooses CF if $\theta_1^* < \theta \leq \theta_2^*$, and chooses CC otherwise.
(ii) The optimal renegotiation-proof fixed prices $\mathbf{b}^\dagger = \left(b_1^{FF\dagger}, b_2^{FF\dagger}, b_2^{CF\dagger}\right)'$ are such that $\underline{b} \equiv b_1^{FF\dagger} = b_2^{FF\dagger} < b_2^{CF\dagger} \equiv \bar{b}$. Moreover, $\underline{b}$, $\bar{b}$, θ_1^* and θ_2^* satisfy

$$\bar{b} = H\left(\theta_2^* - e(\theta_2^*)\right) + \psi\left(e(\theta_2^*)\right), \tag{2.14}$$

$$\underline{b} = \gamma\left[H\left(\theta_1^* - e\left(\theta_1^*\right)\right) + \psi\left(e\left(\theta_1^*\right)\right)\right] + (1-\gamma)\,\bar{b}, \tag{2.15}$$

$$\left(1 - \frac{\alpha}{1+\lambda}\right)\frac{F\left(\theta_2^*\right) - F\left(\theta_1^*\right)}{f\left(\theta_2^*\right)} = \frac{H\left(\theta_2^*\right) - \bar{b}}{H'\left(\theta_2^* - e(\theta_2^*)\right)}. \tag{2.16}$$

Proposition 2.2 above encompasses the results of Proposition 2 in Gagnepain et al. (2013). For the special case of LCC, (2.16) reduces to their Equation (4). Similar to theirs, the intuition for $\underline{b} < \bar{b}$ can be ascribed to that, under renegotiation, fixed prices must be raised sufficiently to induce agents with intermediate efficiency who prefer CR in period-1 to choose FP in period-2.

3 The econometric setting and a test for LCC

For econometric analysis of FP-CR contracts, we keep our focus on the single-period and two-period settings. We aim at two major tasks: to test for the LCC, and to develop identification strategies without the LCC. As to be shown, for both tasks, our methods for the two-period setting are essentially the same as those for the single-period one, with only slight differences which will be detailed later. For this reason, we note the followings: (i) We prioritize to elaborate the latter (i.e. methods for the single-period), only to discuss the former when needing to clarify the differences; (ii)To avoid any confusion, we introduce two (slightly) different sets of notations, for single-period and two-period, respectively. Nevertheless, our prioritization on single-period (as explained in (i)) allows us to keep the notations tidy by sticking with the first set of notations for most parts of the paper.

In this section, we first specify the econometric setting for the rest of the paper. Then we propose a testing procedure for a testable implication of the LCC. In what follows we use upper-case letters for random variables, and lower-case letters for their realizations.

3.1 The econometric setting

From a single-period contract, we, as analysts, observe the realized cost C, the payment Q and a dummy variable D^F that indicates the choice between FF and CR. $D^F = 1$ indicates FF being chosen, and $D^F = 0$ indicates CR. Similarly, from a two-period contract, we observe the per period cost C_t and payment Q_t, for $t = 1, 2$, and dummy variables D^{FF} and D^{CF} that indicate the choice among FF, CF and CC. $D^J = 1$ indicates J being chosen, for $J \in \{FF, CF\}$, and $D^{FF} = D^{CF} = 0$ indicates CC. In addition, we require a binary variable $W \in \{\varpi_1, \varpi_2\}$ that satisfies the following condition, which serves as an exclusion variable:

Condition T.1 *W is independent of θ.*

Depending on whether the empirical application involves single-period contracts or two-period ones, the analyst has at hand either

$$\text{a single-period sample } \left\{c_i, q_i, w_i, d_i^F\right\}_{i=1}^n,$$

i.e., a random sample collected from n observed single-period contracts, or

$$\text{a two-period sample } \left\{c_{it}, q_{it}, w_i, d_i^{FF}, d_i^{CF}\right\}_{i=1,\ldots,n;\, t=1,2},$$

i.e., a random sample collected from n observed two-period contracts.

We introduce a few more notations, most of which are conditional (on W) version of existing ones. Among these notations, the followings are shared by the single-period and the two-period settings: $\psi_j(\cdot) \equiv \psi(\cdot; \varpi_j)$ denotes the disutility function associated with $W = \varpi_j$, and $\psi_j'(e) \equiv \partial\psi(e; \varpi_j)/\partial e$ denotes its derivative; $e(\theta; \varpi_j)$ denotes the optimal effort under FP given θ, conditional on $W = \varpi_j$, which is characterized by the f.o.c.

$$H'(\theta - e(\theta; \varpi_j)) = \psi_j'(e(\theta; \varpi_j)); \tag{3.1}$$

$c(\theta; \varpi_j) \equiv H(\theta - e(\theta; \varpi_j))$ denotes the cost under the optimal effort under FP, conditional on $W = \varpi_j$; For any given $\tau \in [0, 1]$, $\theta(\tau)$ denotes the τ'th quantile of $F(\cdot)$, i.e., the unconditional distribution of θ.

The following notations are exclusive for the single-period setting: θ^{j*} denotes the cut-off value of θ between FP and CR conditional on $W = \varpi_j$. (For more details on θ^{j*}, please refer to Corollary 3.1 (i)); $E(\tau; \varpi_j)$ denotes the τ'th quantile of an effort distribution generated by the transformation $e(\cdot, \varpi_j)$ of the truncated distribution of θ on $[\underline{\theta}, \theta^{j*}]$. In other words, $E(\tau; \varpi_j)$ is the τ'th quantile of the effort conditional on FP being chosen and $W = \varpi_j$; Similarly, $C(\tau; \varpi_j)$ denotes the τ'th quantile of the cost conditional on FP being chosen and $W = \varpi_j$.

The followings notations are exclusive for the two-period setting: θ_t^{j*} denotes the the cut-off value of θ between FP and CR in period t conditional on $W = \varpi_j$, for $j = 1, 2$. (For more details on θ_t^{j*}, please refer to Corollary 3.1 (ii)); $E_t(\tau; \varpi_j)$ denotes the τ'th quantile of an effort distribution generated by the transformation $e(\cdot, \varpi_j)$ of the truncated distribution of θ on $\left[\underline{\theta}, \theta_t^{j*}\right]$. In other words, $E_t(\tau; \varpi_j)$ is the τ'th quantile of the effort in period t conditional on FP being chosen in that period and $W = \varpi_j$; Similarly, $C_t(\tau; \varpi_j)$ denotes the τ'th quantile of the cost in period t conditional on FP being chosen in that period and $W = \varpi_j$.

Regarding the observed variables and their samples, we maintain several regularity conditions throughout the econometric analysis, listed as Assumption 3.1 below.

Assumption 3.1 *The following conditions hold for the single-period setting (or for the two period-setting): (i) $\left\{c_i, q_i, w_i, d_i^F\right\}_{i=1}^n$ (or $\left\{c_{it}, q_{it}, w_i, d_i^{FF}, d_i^{CF}\right\}_{i=1,\ldots,n;\, t=1,2}$) are independent and identically distributed across i for each n; (ii) The density $f(\cdot)$ of θ exists, and is bounded and continuous on a bounded support; (iii) The conditional density $f_{C|W}(c|\varpi_j)$ of cost c under FP (or $f_{C_t|W(c_t|\varpi_j)}$ of period-t cost c_t under FP) for $j = 1, 2$ exist, and are bounded and uniformly continuous in c (or c_t) on bounded supports.*

And we modify conditions in Assumption 2.1 as follows to suit the econometrics analysis:

Assumption 3.2 *(i) $H(\cdot) \geq 0$, $H'(\cdot) > 0$ and $H''(\cdot) > 0$; (ii) $\psi_j(\cdot) \geq 0$, $\psi_j'(\cdot) > 0$, $\psi_j''(\cdot) > 0$ and $\psi_j(0) = 0$, for $j = 1, 2$.*

Under Assumption 3.2, all results in Lemma 2.1, Propositions 2.1 and 2.2 hold conditioned on W. Being perhaps repetitive, we (re)state these results below for convenience of reference later in the paper.

Corollary 3.1 *Let Assumption 3.2 hold. Conditioned on $W = \varpi_j$, it holds in equilibrium that:*
(i) Under FP, the optimal effort $e(\theta; \varpi_j)$ is strictly increasing in θ, with $\partial e(\theta; \varpi_j)/\partial\theta \in (0, 1)$. Moreover, under FP, both $\theta - e(\theta; \varpi_j)$ and $H(\theta - e(\theta; \varpi_j))$ are strictly increasing in θ.
(ii) For the single-period setting, there is a cut-off value $\theta^{j*} \in \left[\underline{\theta}, \bar{\theta}\right]$ such that the agent chooses FP if $\theta \leq \theta^{j*}$, and chooses CR otherwise. And the optimal fixed price b^{j*} and θ^{j*} satisfy

$$b^{j*} = H\left(\theta^{j*} - e\left(\theta^{j*}; \varpi_j\right)\right) + \psi_j\left(e\left(\theta^{j*}; \varpi_j\right)\right), \tag{3.2}$$

$$\left(1 - \frac{\alpha}{1+\lambda}\right) \frac{F\left(\theta^{j*}\right)}{f\left(\theta^{j*}\right)} = \frac{H\left(\theta^{j*}\right) - b^{j*}}{H'\left(\theta^{j*} - e\left(\theta^{j*}; \varpi_j\right)\right)}. \tag{3.3}$$

(iii) For the two-period setting, there are two cut-off values $\theta_1^{j*}, \theta_2^{j*} \in [\underline{\theta}, \bar{\theta}]$ *with* $\theta_1^{j*} < \theta_2^{j*}$ *such that the agent chooses FF if* $\theta \leq \theta_1^{j*}$*, chooses CF if* $\theta_1^{j*} < \theta \leq \theta_2^{j*}$*, and chooses CC otherwise. The optimal renegotiation-proof fixed prices* $\mathbf{b}^{j\dagger} = \left(b_1^{j,FF\dagger}, b_2^{j,FF\dagger}, b_2^{j,CF\dagger}\right)'$ *are such that* $\underline{b}^j \equiv b_1^{j,FF\dagger} = b_2^{j,FF\dagger} < b_2^{j,CF\dagger} \equiv \bar{b}^j$*. Moreover,* $\underline{b}^j$*,* $\bar{b}^j$*,* θ_1^{j*} *and* θ_2^{j*} *satisfy*

$$\bar{b}^j = H\left(\theta_2^{j*} - e\left(\theta_2^{j*}; \varpi_j\right)\right) + \psi_j\left(e\left(\theta_2^{j*}; \varpi_j\right)\right), \tag{3.4}$$

$$\underline{b}^j = \gamma\left[H\left(\theta_1^{j*} - e\left(\theta_1^{j*}; \varpi_j\right)\right) + \psi_j\left(e\left(\theta_1^{j*}; \varpi_j\right)\right)\right] + (1-\gamma)\,\bar{b}^j, \tag{3.5}$$

$$\left(1 - \frac{\alpha}{1+\lambda}\right)\frac{F\left(\theta_2^{j*}\right) - F\left(\theta_1^{j*}\right)}{f\left(\theta_2^{j*}\right)} = \frac{H\left(\theta_2^{j*}\right) - \bar{b}^j}{H'\left(\theta_2^{j*} - e\left(\theta_2^{j*}; \varpi_j\right)\right)}. \tag{3.6}$$

Clearly, Corollary 3.1 (i) - (iii) parallel Lemma 2.1, Propositions 2.1 and 2.2, respectively.

We now establish relationship between certain quantiles, which leads to a testable implication under the LCC, and also turns out to be the key to identify the cost function without the LCC. According to the definitions of $e(\cdot;\varpi_j)$ and $c(\cdot;\varpi_j)$, it holds under FP that $c(\theta;\varpi_j) = H(\theta - e(\theta;\varpi_j))$ for all θ and $j = 1,2$, which immediately implies that $\theta = H^{-1}(c(\theta;\varpi_j)) + e(\theta;\varpi_j)$ for all θ. (Note that the inversibility of $H(\cdot)$ is guaranteed by Assumption 3.2 (i).) Consequently, we have

$$\theta(\tau) = H^{-1}(c(\theta(\tau);\varpi_j)) + e(\theta(\tau);\varpi_j), \forall \tau \in [0,1]. \tag{3.7}$$

In the single-period setting, define $p_j \equiv F\left(\theta^{j*}\right)$, for $j = 1,2$, i.e., the proportion of FP conditioned on $W = \varpi_j$. the τ'th quantile of θ corresponds to the (τ/p_j)'th quantile of the truncated distribution of θ on $\left[\underline{\theta}, \theta^{j*}\right]$ which in turn corresponds to the (τ/p_j)'th quantile of $c(\theta;\varpi_j)$, for $j = 1,2$ and any given $\tau \in [0,p_1] \subseteq [0,p_2]$. Consequently, we can rewrite (3.7) in terms of these quantiles as

$$\theta(\tau) = H^{-1}(C(\tau/p_j;\varpi_j)) + E(\tau/p_j;\varpi_j),\ \forall \tau \in [0,p_j]. \tag{3.8}$$

Similarly, in the two-period setting, it holds that

$$\theta(\tau) = H^{-1}(C_t(\tau/p_{t,j};\varpi_j)) + E_2(\tau/p_{t,j};\varpi_j),\ \forall \tau \in [0,p_{t,j}], \tag{3.9}$$

for $t = 1,2$, where $p_{t,j} \equiv F\left(\theta_t^{j*}\right)$, i.e., the proportion of FP in period t conditioned on $W = \varpi_j$.

3.2 Testing for the LCC

The LCC restricts the cost function to take the form $H(\theta - e) = \beta(\theta - e)$ for some constant coefficient β. To obtain a testable implication of the LCC, we start by noting that, under the LCC, the f.o.c. (3.1) becomes $\psi_j'(e(\theta;\varpi_j)) \equiv \beta$. This together with Assumption 3.2(ii) imply that $e(\theta;\varpi_j) \equiv a_j$ for some positive constant a_j, for all $\theta \in \left[\underline{\theta}, \theta_2^{j*}\right]$ and $j = 1,2$. [5] Consequently,

[5] Since $\psi_1'(\cdot)$ and $\psi_2'(\cdot)$ may differ, it is possible that $a_1 \neq a_2$.

under the LCC, Equation (3.8) becomes $\theta(\tau) = H^{-1}(C(\tau/p_j; \varpi_j)) + a_j$, for all $\tau \in [0,1]$, which in turn implies

$$C(\tau; \varpi_1) - C(\tau p; \varpi_2) \equiv \text{a constant}\,(= \beta(a_2 - a_1))\,,\ \forall \tau \in [0,1], \tag{3.10}$$

with $p \equiv p_1/p_2$. Next, we show that (3.10) is testable. Specifically, we show how to test

$$\begin{aligned} \mathbb{H}_0: &\quad C(\tau; \varpi_1) - C(\tau p; \varpi_2) \equiv \text{a constant, for } \tau \in [0,1] \text{ almost sure };\\ \mathbb{H}_1: &\quad \text{Otherwise,} \end{aligned}$$

focusing on the single-period setting. The testing procedure to be presented only needs a few simple modifications to work for the two-period setting, as we elaborate at the end of this subsection.

The Test Statistic

Based on a single-period sample $\{c_i, w_i, d_i^F\}_{i=1}^n$, we construct the following *Cramér-von Mises* (CvM) type test statistic

$$T_n \equiv \int_0^1 n_f \left\{ \left[\widehat{C}(\tau; \varpi_1) - \widehat{C}(\tau\hat{p}; \varpi_2)\right] - \left[\widehat{C}(\nu; \varpi_1) - \widehat{C}(\nu\hat{p}; \varpi_2)\right] \right\}^2 d\tau \tag{3.11}$$

where: (i) $\nu \in (0,1)$ is a pre-selected constant that serves as the reference quantile level; (ii) $n_f \equiv \sum_{i=1}^n d_i^F$ is the number of FP contracts in the sample. (iii) $\hat{p} = \hat{p}_1/\hat{p}_2$ is an estimator for $p = p_1/p_2$, with

$$\hat{p}_j \equiv \sum_{i=1}^n d_i^F \mathbf{1}(w_i = \varpi_j) / \sum_{i=1}^n \mathbf{1}(w_i = \varpi_j), \text{ for } j = 1, 2; \tag{3.12}$$

(iv) $\left\{\widehat{C}(\tau, \varpi_1), \widehat{C}(\tau\hat{p}, \varpi_2)\right\}$ are obtained from quantile regressions based on the FP subsample $\left\{c_{f_i}, w_{f_i}, d_{f_i}^F\right\}_{i=1}^{n_f}$. Specifically, $\widehat{C}(\tau, \varpi_1) = \widehat{q}_1(\tau)$ and $\widehat{C}(\tau\hat{p}; \varpi_2) = \widehat{q}_1(\tau\hat{p}) + \widehat{q}_2(\tau\hat{p})$, where

$$\begin{aligned} \{\widehat{q}_1(\tau), \widehat{q}_2(\tau)\} &= \underset{q_1,\, q_2}{\operatorname{argmin}} \sum_{i=1}^{n_f} \rho_\tau\left(c_{f_i} - q_1 - q_2 \cdot \mathbf{1}(w_{f_i} = \varpi_2)\right),\\ \{\widehat{q}_1(\tau\hat{p}), \widehat{q}_2(\tau\hat{p})\} &= \underset{q_1,\, q_2}{\operatorname{argmin}} \sum_{i=1}^{n_f} \rho_{\tau\hat{p}}\left(c_{f_i} - q_1 - q_2 \cdot \mathbf{1}(w_{f_i} = \varpi_2)\right), \end{aligned}$$

with $\rho_a(b) \equiv b[a - \mathbf{1}(b < 0)]$ for any given $\{a, b\}$.

Based on a uniform weak convergence result regarding a corresponding *quantile regression process* (indexed by τ) by Angrist et al. (2006) (detailed Lemma A.3 in Appendix A and its proof), we establish the asymptotic behavior of T_n, as follows

Theorem 1 (Asymptotic behavior of T_n) *Let Assumptions 3.1, 3.2, and Condition T.1 hold.*
(i) If $\mathbb{H}_0$ holds true, then

$$T_n \quad \xrightarrow{\mathcal{L}} \quad \mathcal{F}\left(G\left(\cdot\right),\, N\left(0, \sigma_p^2\right)\right)$$

for some functional $\mathcal{F}$, where $G(\cdot)$ is a tight Gaussian process on $L^{\infty}[0,1]$ and $\sigma_p^2 < \infty$;
(ii) Under any fixed alternative, $T_n \xrightarrow{p} +\infty$.

Theorem 1 shows that T_n converges in distribution to a tight distribution under the null, and diverges to ∞ in probability under any fixed alternative.

Bootstrap implementation

The null asymptotic distribution is unfamiliar. We adopt a bootstrap procedure to obtain critical values for the test. The justification of its asymptotic validity is standard. Specifically, based on a bootstrap sample $\left\{c_i^*, w_i^*, d_i^{F*}\right\}_{i=1}^n$, we construct the the bootstrap statistic

$$T_n^* \quad = \quad \int_0^1 n_f^* \cdot V^*\left(\tau\right)^2 d\tau, \tag{3.13}$$

with a re-centered term

$$\begin{aligned} V^*\left(\tau\right) \quad \equiv \quad & \left\{\left[\widehat{C}^*\left(\tau;\varpi_1\right) - \widehat{C}^*\left(\tau\hat{p}^*;\varpi_2\right)\right] - \left[\widehat{C}^*\left(\nu;\varpi_1\right) - \widehat{C}^*\left(\nu\hat{p}^*;\varpi_2\right)\right]\right\} \\ & - \left\{\left[\widehat{C}\left(\tau;\varpi_1\right) - \widehat{C}\left(\tau\hat{p};\varpi_2\right)\right] - \left[\widehat{C}\left(\nu;\varpi_1\right) - \widehat{C}\left(\nu\hat{p};\varpi_2\right)\right]\right\}, \end{aligned}$$

where: (i) $n_f^* \equiv \sum_{i=1}^n \mathbf{1}\left\{d_i^{F*} = 1\right\}$; (ii) $\hat{p}^* = \hat{p}_1^*/\hat{p}_2^*$ with

$$\hat{p}_j^* \equiv \sum_{i=1}^n d_i^{F*}\mathbf{1}\left(w_i^* = \varpi_j\right) / \sum_{i=1}^n \mathbf{1}\left(w_i^* = \varpi_j\right);$$

(iii)$\widehat{C}^*\left(\tau,\varpi_1\right) = \widehat{q}_1^*\left(\tau\right)$ and $\widehat{C}^*\left(\tau\hat{p}^*;\varpi_2\right) = \widehat{q}_1^*\left(\tau\hat{p}^*\right) + \widehat{q}_2^*\left(\tau\hat{p}^*\right)$, with

$$\begin{aligned} \left\{\widehat{q}_1^*\left(\tau\right), \widehat{q}_2^*\left(\tau\right)\right\} \quad &= \quad \underset{q_1,\, q_2}{\operatorname{argmin}} \sum_{i=1}^{n_f^*} \rho_\tau\left(c_{f_i}^* - q_1 - q_2 \cdot \mathbf{1}\left(w_{f_i}^* = \varpi_2\right)\right) \\ \left\{\widehat{q}_1^*\left(\tau\hat{p}^*\right), \widehat{q}_2^*\left(\tau\hat{p}^*\right)\right\} \quad &= \quad \underset{q_1,\, q_2}{\operatorname{argmin}} \sum_{i=1}^{n_f^*} \rho_{\tau\hat{p}^*}\left(c_{f_i}^* - q_1 - q_2 \cdot \mathbf{1}\left(w_{f_i}^* = \varpi_2\right)\right), \end{aligned}$$

based on $\left\{c_{f_i}^*, w_{f_i}^*, d_{f_i}^{F*}\right\}_{i=1}^{n_f^*}$, i.e., the FP subsample of $\left\{c_i^*, w_i^*, d_i^{F*}\right\}_{i=1}^n$.

Test in the two-period setting: needed modifications

Based on the second period observations $\left\{c_{i2}, q_{i2}, w_i, d_i^{FF}, d_i^{CF}\right\}_{i=1}^n$, the testing procedure needs a few modifications to work for the two-period setting.[6] Let $\left\{c_{f_i,2}, q_{f_i,2}, w_{f_i}, d_{f_i}^{FF}, d_{f_i}^{CF}\right\}_{i=1}^{\tilde{n}_f}$ be the subsample of contracts with FP in the second period (i.e. FF and CF contracts, but no CC ones), with $\tilde{n}_f = \sum_{i=1}^n \mathbf{1}\left(d_i^{FF} + d_i^{CF} = 1\right)$. In addition, let $\left\{c_{i2}^*, q_{i2}^*, w_i^*, d_i^{FF*}, d_i^{CF*}\right\}_{i=1}^n$ and $\left\{c_{f_i,2}^*, q_{f_i,2}^*, w_{f_i}^*, d_{f_i}^{FF*}, d_{f_i}^{CF*}\right\}_{i=1}^{\tilde{n}_f^*}$ be the corresponding bootstrap sample and FP subsample, with $\tilde{n}_f^* \equiv \sum_{i=1}^n \mathbf{1}\left(d_i^{FF*} + d_i^{CF*} = 1\right)$. The modifications needed are as follows: (i) Replace n_f and n_f^* by $\tilde{n}_f$ and $\tilde{n}_f^*$, respectively; (ii) Replace c_{f_i} and $c_{f_i}^*$ by $c_{f_i,2}$ and $c_{f_i,2}^*$, respectively; (iii) For $j = 1, 2$, replace $\hat{p}_j$ and $\hat{p}_j^*$ by $\hat{p}_{2,j}$ and $\hat{p}_{2,j}^*$, respectively, which are defined as

$$\hat{p}_{2,j} = \frac{\sum_{i=1}^n \left(d_i^{FF} + d_i^{CF}\right) \mathbf{1}\left(w_i = \varpi_j\right)}{\sum_{i=1}^n \mathbf{1}\left(w_i = \varpi_j\right)}, \quad \hat{p}_{2,j}^* = \frac{\sum_{i=1}^n \left(d_i^{FF*} + d_i^{CF*}\right) \mathbf{1}\left(w_i^* = \varpi_j\right)}{\sum_{i=1}^n \mathbf{1}\left(w_i^* = \varpi_j\right)}. \tag{3.14}$$

4 Identification

Note that the single-period and two-period settings share almost all model primitives $\mathcal{S} \equiv [F(\cdot), H(\cdot), \psi_1(\cdot), \psi_2(\cdot), \alpha/(1+\lambda)]$, except for γ which measures time preference and is featured only in the latter setting. Also note that α and λ are not separately identifiable.[7] Therefore, we treat $\alpha/(1+\lambda)$ as a whole for identification.

In this section we develop regularity conditions for identification of $\mathcal{S}$ under the single-period setting, and $\mathcal{S}$ and γ under the two-period one. As to be shown, non of these conditions requires the LCC to hold. Besides the observable variables specified in Section 3.1 (i.e., $\left\{C, Q, WD^F\right\}$ under the single-period setting, or $\left\{C_t, Q_t, W, D^{FF}, D^{CF}\right\}_{t=1,2}$ under the two-period setting), we may also observe some other characteristics of the principal or/and the agent, denoted by a vector of covariates Z. Though, we make the following arrangements for elaborating our identification strategies: (i) We suppress Z for simplicity of notations, noting that every step involved in the identification strategies would work just fine when conditioned on Z; (ii) We still prioritize the single-period setting, which is our default setting unless otherwise specified.

The following lemma shows that $\mathcal{S}$ is not point identified without further restrictions.

Lemma 4.1 ***(Observational equivalence)*** *Suppose $\mathcal{S}$ satisfies Assumption 3.2. Then, $\mathcal{S} \equiv [F, H, \psi_1, \psi_2, \alpha/(1+\lambda)]$ is observationally equivalent to $\widetilde{\mathcal{S}} \equiv [\widetilde{F}, \widetilde{H}, \widetilde{\psi}_1, \widetilde{\psi}_2, \alpha/(1+\lambda)]$ with $\widetilde{F}(\cdot) = F(\cdot/\xi)$, $\widetilde{H}(\cdot) = H(\cdot/\xi)$, and $\widetilde{\psi}_j(\cdot) = \psi_j(\cdot/\xi)$ for $j = 1, 2$, for any given constant $\xi > 0$.*

To achieve identification, we normalize $\underline{\theta}$ and $\underline{e}$ (i.e., lower bounds of the supports for θ and e, respectively,) to be known[8].

[6] Recall that, under the two-period setting, a sample takes the form $\left\{c_{it}, q_{it}, w_i, d_i^{FF}, d_i^{CF}\right\}_{i=1,\ldots,n;\, t=1,2}$.

[7] To see this, note that the economic model remain the same for difference pairs of $\{\alpha, \lambda\}$ as long as the ratio $\alpha/(1+\lambda)$ is the same.

[8] Similar normalization conditions are imposed in Perrigne and Vuong (2011, 2012) for identification.

4.1 Recovering the optimal effort

First, we identify the distribution of E under FP by adopting Schennach and Hu (2013)'s method, for which we require two effort-related proxies that satisfy the following assumption:

Assumption 4.1 *There exist two effort-related proxy variables X and Y such that*

$$X = E + V_1 \quad \text{and} \quad Y = m(E) + V_2,$$

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

According to Assumption 4.1, X can be regarded as an observed measure of E with measurement errors. The relation between Y and E is more flexible as we do not restrict the functional form of $m(\cdot)$. For example, agent's effort-related performance is a potential candidate for Y. See Cicala (2015) for a discussion on the plausibility of employing cost-related variables to infer the agent's effort. In our empirical study of public transportation service contracts in France, X and Y are picked to be the share of drivers among all the employees (which primarily consist of drivers and engineers) and the labor fee, respectively. We discuss the justification of such a proxies choice in the empirical section. Note that Assumption 4.1 is less restrictive then requiring two direct measurements of a latent variable, which is imposed by many existing papers to identify various structural models. For instance, Li (2002) requires $m(\cdot)$ to be an identity function.

Denote by $F_{E|W}(\cdot|\varpi_j)$ the conditional CDF of effort (under FP) on $W = \varpi_j$. It follows from Theorem 1 of Schennach and Hu (2013) that $m(\cdot)$ and$F_{E|W}(\cdot|\varpi_j)$ are nonparametrically identifiable from the conditional distribution of (X, Y) on $W = \varpi_j$, for $j = 1, 2$, except for some rather specific data generating processes (DGP), which impose little restrictions to our model, as discussed in Appendix.

Once $F_{E|W}(\cdot|\varpi_j)$ is identified, we can recover the unobserved effort corresponding to an observe cost c under FP conditioned on $W = \varpi_j$ according to Corollary 3.1: (i) c corresponds to some $\theta \in \left[\underline{\theta}, \theta^{j*}\right]$ such that $c = c(\theta; \varpi_j) \equiv H(\theta - e(\theta; \varpi_j))$; (ii) it necessarily holds that $c \in \left[c(\underline{\theta}; \varpi_j), c\left(\theta^{j*}; \varpi_j\right)\right] \equiv \left[\underline{c}^j, \bar{c}^j\right]$; (iii) the mapping from $c = c(\theta; \varpi_j)$ to its corresponding effort $e = e(\theta; \varpi_j)$ is strictly increasing and bijective from $\left[\underline{c}^j, \bar{c}^j\right]$ to $\left[\underline{e}^j, \bar{e}^j\right] \equiv \left[e(\underline{\theta}; \varpi_j), e\left(\theta^{j*}; \varpi_j\right)\right]$. Consequently, although unobserved, $e = e(\theta; \varpi_j)$ can be recovered as

$$e = F_{E|W}^{-1}(F_{C|W}(c|\varpi_j)|\varpi_j), \; \forall\, c \in [\underline{c}^j, \bar{c}^j], \tag{4.1}$$

where $F_{C|W}(\cdot|\varpi_j)$ is the conditional CDF of cost (under FP) on $W = \varpi_j$. Under CR, the effort level is simply zero. We summarize this identification result in the following proposition:

Proposition 4.1 *Let Assumptions 3.1, 3.2 and 4.1 hold. Conditioned on W, the (optimal) effort level associated with any observed cost c is identifiable.*

4.2 Identification of the cost function $H(\cdot)$

To identify $H(\cdot)$, we exploit Equation (3.8): $\theta(\tau) = H^{-1}(C(\tau/p_j;\varpi_j)) + E(\tau/p_j;\varpi_j)$, $\forall \tau \in [0, p_j]$, which in turn implies that

$$H^{-1}(C(\tau/p_1;\varpi_1)) = H^{-1}(C(\tau/p_2;\varpi_2)) + \Delta\tilde{E}(\tau),\ \forall\, \tau \in [0, \min\{p_1, p_2\}] \tag{4.2}$$

with $p_j \equiv F(\theta^{j*})$ as previously defined, and $\Delta\tilde{E}(\tau) \equiv E(\tau/p_2;\varpi_2) - E(\tau/p_1;\varpi_1)$. Although Condition T.1 suffices for (3.8) and (4.2) to hold, further requirements on W are needed for identification, which are listed in Assumption 4.2.

Assumption 4.2 *The binary variable W satisfies: (i) W is independent of θ; (ii) The disutility from effort is dependent on W such that $\psi_1'(e) \geq \psi_2'(e)$ for all $e \geq \underline{e}$, with equality holding only at $e = \underline{e}$, i.e., $\psi_1'(\underline{e}) = \psi_2'(\underline{e})$.*

Assumption 4.2 (i) repeats Condition T.1. Assumption 4.2 (ii) specifies a further requirement: agents with $W = \varpi_1$ incur more marginal disutility than those with $W = \varpi_2$ does, and the two marginal curve $\{\psi_1'(\cdot), \psi_2'(\cdot)\}$ cross only once, at the lower bound $\underline{e}$. Similar single-crossing or finite-crossing conditions are widely imposed for identification purpose in the literature, see, e.g. Chesher (2003), Chernozhukov and Hansen (2005), Heckman et al. (2010), and Torgovitsky (2015), among others.

Assumptions 3.2 and 4.2, in conjunction with the f.o.c. $H'(\underline{\theta} - e(\underline{\theta};\varpi_j)) = \psi_j'(e(\underline{\theta};\varpi_j))$, imply the followings: (i) $e(\underline{\theta};\varpi_1) = e(\underline{\theta};\varpi_2) = \underline{e}$, hence $\psi_1'(e(\underline{\theta};\varpi_1)) = \psi_2'(e(\underline{\theta};\varpi_2))$. Similarly, $\underline{c}^j = H(\underline{\theta} - e(\underline{\theta};\varpi_j)) \equiv H(\underline{\theta} - \underline{e}) = \underline{c}$ for $j = 1, 2$; (ii) $e(\theta;\varpi_1) < e(\theta;\varpi_2)$ for any $\theta > \underline{\theta}$; (iii) Regarding θ^{j*}, i.e., the cut-off type conditioned $W = \varpi_j$, we have $\theta^{1*} \leq \theta^{2*}$. Consequently, $p_1 \leq p_2$. (Similarly, for the two-period setting, we have $\theta_t^{1*} \leq \theta_t^{2*}$ and $p_{t,1} \leq p_{t,2}$ for $t = 1, 2$.)

A key to identifying $H(\cdot)$ is to establish for the term $\Delta\tilde{E}(\tau)$ in (4.2) that

$$\Delta\tilde{E}(\tau) > 0 \text{ for all } \tau \in (0, p_1], \tag{4.3}$$

which we show in the proof of Proposition 4.2 in Appendix A. Also recall that, for $j = 1, 2$, $F_{C|W}(\cdot|\varpi_j)$ represents the cost distribution under FP conditional on $W = \varpi_j$, hence with support $[\underline{c}, \bar{c}^j]$. So, for any $c \in [\underline{c}, \bar{c}^1]$, there exists a unique quantile level $\tau_0(c) \in [0, 1]$ such that $c = C(\tau_0(c);\varpi_1)$.[9] If $c = \underline{c}$, it holds, according to Corollary 3.1 (i), that $\tau_0(c) = 0$ and that $H^{-1}(\underline{c})$ is immediately identified as $H^{-1}(\underline{c}) = \underline{\theta} - \underline{e}$. If $c \in (\underline{c}, \bar{c}^1]$, it holds, again according to Corollary 3.1 (i), that $\tau_0(c) \in (0, 1]$. Moreover, for $c \in (\underline{c}, \bar{c}^1]$, it follows from (4.2) and (4.3) that $C(\tau_0(c);\varpi_1) = C(\tau_0(c) \cdot p_1/p_2;\varpi_2) + \Delta\tilde{E}(\tau_0(c)p_1) > C(\tau_0(c) \cdot p_1/p_2;\varpi_2) > \underline{c}$. It follows from the same logic that there exists $\tau_1(c) \in (0, \tau_0(c))$ such that $C(\tau_0(c) \cdot p_1/p_2;\varpi_2) = C(\tau_1(c);\varpi_1) > C(\tau_1(c) \cdot p_1/p_2;\varpi_2)$. Recursively, for any $c \in (\underline{c}, \bar{c}^1]$, we can establish a strictly decreasing sequence of quantiles $\{\tau_k(c)\}_{k=0}^{\infty}$ such that $C(\tau_k(c).p_1/p_2;\varpi_2) = C(\tau_{k+1}(c);\varpi_1) >$

[9] We implicitly assume $F_{C|W}(\cdot|\varpi_j)$ to be continuous and strictly increasing on its support.

$C\left(\tau_{k+1}(c).p_1/p_2;\varpi_2\right)$, with a initial condition that $C\left(\tau_0(c);\varpi_1\right)=c$. Iterating (4.2) $m+1$ times according to $\{\tau_k(c)\}_{k=0}^{\infty}$ yields:

$$\sum\nolimits_{k=0}^{m}\Delta\tilde{E}(\tau_k(c).p_1)=H^{-1}(c)-H^{-1}\left(C\left(\tau_m(c).p_1/p_2;\varpi_2\right)\right).$$

In the proof of Proposition 4.2 (in Appendix A), we show the followings: (i) The decreasing sequence $\{\tau_k\left(c\right)\}_{k=0}^{\infty}$ converges to 0; (ii) $\sum_{k=0}^{\infty}\Delta\tilde{E}(\tau_k(t).p_1)<\infty$. Consequently, taking limit of $m\to\infty$ over the equation above and rearranging terms yields

$$H^{-1}\left(c\right)=H^{-1}\left(C\left(0;\varpi_2\right)\right)+\sum\nolimits_{k=0}^{\infty}\Delta\tilde{E}(\tau_k(c).p_1)=\underline{\theta}-\underline{e}+\sum\nolimits_{k=0}^{\infty}\Delta\tilde{E}(\tau_k(c).p_1) \tag{4.4}$$

for any $c\in\left(\underline{c},\bar{c}^1\right]$. Note that the whole sequence $\{\tau_k(c)\}_{k=0}^{\infty}$ is identifiable, because $F_{C|W}\left(\cdot|\varpi_j\right)$ are directly identifiable as both C and W are observed. And $\Delta\tilde{E}\left(\tau_k\cdot p_1\right)=E\left(\tau_k.p_1/p_2;\varpi_2\right)-E\left(\tau_k;\varpi_1\right)$ is identifiable by Proposition 4.1. Therefore, (4.4) identifies $H^{-1}\left(c\right)$ for any $c\in\left(\underline{c},\bar{c}^1\right]$. Note that Guerre et al. (2009) and D'Haultfœuille and Février (2020) achieve identification by exploit quantile relations that are similar to (4.4). We formalize the identification result for $H\left(\cdot\right)$ in the following proposition:

Proposition 4.2 *Let Assumptions 3.1, 3.2, 4.1 and 4.2 hold. $H^{-1}(\cdot)$ is nonparametrically identified on $[\underline{c},\bar{c}^1]$ as*

$$H^{-1}(c)=\begin{cases}\underline{\theta}-\underline{e}, & \text{for } c=\underline{c},\\ \underline{\theta}-\underline{e}+\sum_{k=0}^{\infty}\Delta\tilde{E}(\tau_k(c)\cdot p_1), & \text{for } c\in\left(\underline{c},\bar{c}^1\right].\end{cases} \tag{4.5}$$

As shown in the proof of Proposition 4.2, the single-crossing condition of Assumption 4.2 (ii) plays a crucial role: It helps to locate the limiting pointing of the sequence $\{\tau_k\left(c\right)\}_{k=0}^{\infty}$, which eventually leads to the identification result above. Nevertheless, it is worth noting that whether the crossing happens at a boundary point ($\underline{e}$ as in Assumption 4.2 (ii)) or not is nonessential, in the sense that $H\left(\cdot\right)$ can be identified (on certain range) by similar iterative steps when Assumption 4.2 (ii)'s single-crossing at the boundary is altered to single-crossing at an interior point. We formalize this alternative identification result in Appendix C, labeled as Corollary C.1.

It also worth noting that, with no additional assumptions, it is impossibly to nonparametrically identify $H(\cdot)$ on $(\theta_2^{1*}-\bar{e}_2^1,\overline{\theta}]$. And the reason is as follows: Agents with $\theta\in(\theta_2^{1*},\overline{\theta}]$ would choose CR, which provides very limited information on $H(\cdot)$. With θ being unobserved and $e\equiv 0$, all we learn from CR contracts are merely that $c=H(\theta)$.

4.3 Identification of other model primitives

Based on identification of $H^{-1}\left(\cdot\right)$ on $\left[\underline{c},\bar{c}^1\right]$, we proceed to identify other model primitives, in the following order: the type distribution, the disutility functions, and the ratior $\alpha/\left(1+\lambda\right)$.

The type distribution

Conditioned on $W = \varpi_1$, any realized cost $c \in \left[\underline{c}, \bar{c}^1\right]$ is necessarily associated with the FP option, with a corresponding $\theta \in \left[\underline{\theta}, \theta^{1*}\right]$. And any $c > \bar{c}^1$ is necessarily associated with CR, with a corresponding $\theta \in \left(\theta^{1*}, \bar{\theta}\,\right]$. More specifically, conditioned on $W = \varpi_1$, the corresponding θ for any c satisfies

$$\theta = \begin{cases} H^{-1}(c) + F^{-1}_{E|W}\left(F_{C|W}\left(c|\varpi_1\right)|\varpi_1\right), \text{ for } c \in \left[\underline{c}, \bar{c}^1\right]; \\ H^{-1}(c), \text{ for } c \in (\bar{c}^1, \bar{c}]. \end{cases} \tag{4.6}$$

Given the identification results we have already established, we can recover the corresponding θ for any observed value of cost $c \in \left[\underline{c}, \bar{c}^1\right]$ by (4.6). Consequently, we can identify the truncated distribution of θ on $\left[\underline{\theta}, \theta^{1*}\right]$ (i.e., conditional on $\theta \in \left[\underline{\theta}, \theta^{1*}\right]$), for which we denote by $G(\cdot)$ and $g(\cdot)$ its CDF and pdf. We can then identify the (unconditional) CDF and pdf of θ on $\left[\underline{\theta}, \theta^{1*}\right]$ as

$$F(\cdot) = G(\cdot)F\left(\theta^{1*}\right), \quad f(\cdot) = g(\cdot)F\left(\theta^{1*}\right), \tag{4.7}$$

noting that $F\left(\theta^{1*}\right)$ is readily identifiable as $F\left(\theta^{1*}\right) = \mathbb{E}\left(D^F|W = \varpi_1\right)$. However, without additional assumptions, it is impossibly to identify $F(\cdot)$ or $f(\cdot)$ on the rest of their support, i.e., $\left(\theta^{1*}, \bar{\theta}\,\right]$, as $H^{-1}(\cdot)$ is not identified on $\left(\bar{c}^1, \bar{c}\right]$.

The disutility functions

To identify $\psi_1(\cdot)$ and $\psi_2(\cdot)$, recall (4.1) which characterizes the effort e corresponding to any $c \in \left[\underline{c}, \bar{c}^1\right]$, conditioned on $W = \varpi_j$, as $e = F^{-1}_{E|W}\left(F_{C|W}\left(c|\varpi_j\right)|\varpi_j\right)$. Inversely, conditioned on $W = \varpi_j$, the cost c corresponding to any effort $e \in \left[\underline{e}, \bar{e}^j\right]$ satisfies

$$c = F^{-1}_{C|W}\left(F_{E|W}\left(e|\varpi_j\right)|\varpi_j\right) \in \left[\underline{c}, \bar{c}^j\right],$$

which, together with the f.o.c. (3.1), imply that

$$\psi_j'(e) = H'\left(H^{-1}\left(F^{-1}_{C|W}\left(F_{E|W}\left(e|\varpi_j\right)|\varpi_j\right)\right)\right), \forall e \in \left[\underline{e}, \bar{e}^j\right]. \tag{4.8}$$

A further restriction on $\psi_j(\cdot)$ can be obtained from (3.2): $\psi_j\left(\bar{e}^j\right) = b^{j*} - \bar{c}^j$, which, together with (4.8), identify $\psi_j(\cdot)$ as

$$\psi_j(e) = b^{j*} - \bar{c}^j - \int_e^{\bar{e}^j} H'\left(H^{-1}\left(F^{-1}_{C|W}\left(F_{E|W}\left(v|\varpi_j\right)|\varpi_j\right)\right)\right) dv, \ \forall e \in \left[\underline{e}, \bar{e}^j\right]. \tag{4.9}$$

The ratio $\alpha/(1+\lambda)$

Rearranging terms for Eq (3.3) yields

$$\begin{aligned} \frac{\alpha}{1+\lambda} &= 1 - \frac{H\left(\theta^{1*}\right) - b^{1*}}{H'\left(\theta^{1*} - e\left(\theta^{1*};\varpi_1\right)\right)} \frac{f\left(\theta^{1*}\right)}{F\left(\theta^{1*}\right)} \\ &= 1 - \frac{\underline{c}^1_{CR} - b^{1*}}{H'\left(H^{-1}\left(\bar{c}^1\right)\right)} \frac{f\left(\theta^{1*}\right)}{\mathbb{E}\left(D^F|W = \varpi_1\right)}, \end{aligned} \tag{4.10}$$

with $\underline{c}_{CR}^{1} \equiv H\left(\theta^{1*}\right)\left(> H\left(\theta^{1*}-\bar{e}^{1}\right)=\bar{c}^{1}\right)$ being the lower bound for cost under CR, conditioned on $W=\varpi_1$. Since all terms of (4.10) are either directly identifiable from the observed variables or already shown identifiable, (4.10) identifies $\alpha/(1+\lambda)$.

4.4 Identification for the two-period setting

Having established identification results for the single-period setting in Sections 4.1 - 4.3, now we show briefly similar results holding for the two-period setting.

Like for the single-period setting, identification for the two-period setting starts with recovering the observed cost, for which proxies X_t and Y_t that satisfy the following conditions are required:

$$X_t = E_t + V_{t,1} \quad \text{and} \quad Y_t = m_t(E_t) + V_{t,2}, \text{ for } t=1,2,$$

for unknown functions $m_t(\cdot)$, where E_t, $V_{t,1}$, and $V_{t,2}$ are mutually independent with $\mathbb{E}\left(V_{t,1}\right) = \mathbb{E}\left(V_{t,2}\right)=0$. According to Corollary 3.1 (iii) it necessarily holds for any cost under FP in period-t, c_t, that

$$c_t \in \left[c\left(\underline{\theta};\varpi_j\right), c\left(\theta_t^{j*};\varpi_j\right)\right] \equiv \left[\underline{c}, \bar{c}_t^{j}\right], \text{ for } t=1,2.$$

Moreover, the mapping from $c_t = c\left(\theta;\varpi_j\right)$ to its corresponding effort $e_t = e\left(\theta;\varpi_j\right)$ is strictly increasing and bijective from $\left[\underline{c},\bar{c}_t^{j}\right]$ to $\left[\underline{e},\bar{e}_t^{j}\right] \equiv \left[e\left(\underline{\theta};\varpi_j\right), e\left(\theta_t^{j*};\varpi_j\right)\right]$. Therefore, although $e_t = e\left(\theta;\varpi_j\right)$ is unobserved, it can be recovered as

$$e_t = F_{E_t|W}^{-1}\left(F_{C_t|W}\left(c_t|\varpi_j\right)|\varpi_j\right), \; \forall\, c_t \in [\underline{c},\bar{c}_t^{j}], \tag{4.11}$$

where $F_{C_t|W}\left(\cdot|\varpi_j\right)$ is the conditional CDF of period-t cost (under FP) on $W=\varpi_j$, whose support is $[\underline{c},\bar{c}_t^{j}]$. And $F_{E_t|W}\left(\cdot|\varpi_j\right)$ is the conditional CDF of period-t effort (under FP) on $W=\varpi_j$, whose support is $[\underline{e},\bar{e}_t^{j}]$.

The cost function and type distribution

To identify $H(\cdot)$ and the distribution of θ, we adopt the same identification strategies as in Section 4.2 and 4.3, respectively, using observations from either one of the two periods. Using period-t observation, we are able to identify $H^{-1}(\cdot)$ on $[\underline{c},\bar{c}_t^{1}]$ as

$$H^{-1}(c) = \begin{cases} \underline{\theta}-\underline{e}, & \text{for } c=\underline{c}, \\ \underline{\theta}-\underline{e}+\sum_{k=0}^{\infty}\Delta\tilde{E}_t\left(\tau_k(c)\cdot p_{t,1}\right), & \text{for } c\in\left(\underline{c},\bar{c}_t^{1}\right], \end{cases} \tag{4.12}$$

where $\Delta\tilde{E}_t(\tau) \equiv E_t\left(\tau/p_{t,2};\varpi_2\right) - E_t\left(\tau/p_{t,1};\varpi_1\right)$ for any $\tau\in[0,p_{t,1}]$, and $p_{t,j} = F\left(\theta_t^{j*}\right) = \mathbb{E}\left(D^{FF}|W=\varpi_j\right)$ as in (3.9). Noting that, conditioned on $W=\varpi_1$, any $c_t\in\left[\underline{c},\bar{c}_t^{1}\right]$ is necessarily associated with FP in period-t, with a corresponding $\theta\in\left[\underline{\theta},\theta_t^{1*}\right]$. Following the same reasoning underlying (4.6), we can recover the corresponding θ for any period-t cost $c_t \in \left[\underline{c}, \bar{c}_t^1\right]$ conditioned on $W = \varpi_1$ as

$$\theta = H^{-1}(c_t) + F_{E_t|W}^{-1}\left(F_{C_t|W}\left(c_t|\varpi_1\right)|\varpi_1\right), \forall c \in \left[\underline{c}, \bar{c}_t^1\right]. \tag{4.13}$$

Consequently, we can identify $G(\cdot)$ and $g(\cdot)$, the truncated CDF and pdf of θ on $\left[\underline{\theta}, \theta_t^{1*}\right]$. We can then identify the (unconditional) CDF and pdf of θ on $\left[\underline{\theta}, \theta_t^{1*}\right]$ as $F(\cdot) = G(\cdot)F\left(\theta_t^{1*}\right)$ and $f(\cdot) = g(\cdot)F\left(\theta_t^{1*}\right)$, respectively, noting that $F\left(\theta^{1*}\right)$ is readily identifiable as

$$F\left(\theta_t^{1*}\right) = \begin{cases} \mathbb{E}\left(D^{FF}|W = \varpi_1\right), \text{ for } t = 1, \\ \mathbb{E}\left(D^{FF} + D^{CF}|W = \varpi_1\right), \text{ for } t = 2. \end{cases}$$

To summarize, based on period-t observations (for $t = 1$ or 2), $H(\cdot)$ and $F(\cdot)$ (hence also $f(\cdot)$) are identified on $[\underline{c}, \bar{c}_t^1]$ and $\left[\underline{\theta}, \theta_t^{1*}\right]$, respectively. Recall from Corollary 3.1 (i) and (iii) that $\theta_2^{1*} > \theta_1^{1*}$ and $\bar{c}_2^1 > \bar{c}_1^1$. This implies that, for the purpose of identifying $H(\cdot)$ and $F(\cdot)$, observations from period-2 are more informative than those from period-1, in the sense that period-2 observations identify both functions on larger intervals (i.e., $[\underline{c}, \bar{c}_2^1] \supset [\underline{c}, \bar{c}_1^1]$ and $[\underline{\theta}, \bar{\theta}_2^1] \supset [\underline{\theta}, \bar{\theta}_1^1]$).

The disutility functions

Similar to (4.8), identification of $\psi_1(\cdot)$ and $\psi_2(\cdot)$ is based on the differential equation

$$\psi_j'(e_t) = H'\left(H^{-1}\left(F_{C_t|W}^{-1}\left(F_{E_t|W}\left(e_t|\varpi_j\right)|\varpi_j\right)\right)\right), \forall e_t \in \left[\underline{e}, \bar{e}_t^j\right],$$

which, together with the condition $\psi_j\left(\bar{e}_2^j\right) = \bar{b}^j - H\left(\theta_2^{j*} - \bar{e}_2^j\right) = \bar{b}^j - \bar{c}_2^j$ implied by (3.4), identify $\psi_j(\cdot)$ as

$$\psi_j(e) = \bar{b}^j - \bar{c}_2^j - \int_e^{\bar{e}_2^j} H'\left(H^{-1}\left(F_{C_t|W}^{-1}\left(F_{E_t|W}\left(v|\varpi_j\right)|\varpi_j\right)\right)\right) dv, \ \forall e \in \left[\underline{e}, \bar{e}_2^j\right]. \tag{4.14}$$

The ratio $\alpha/(1+\lambda)$

In the two-period setting, identification of $\alpha/(1+\lambda)$ requires using observations from both periods, and is based on (3.6) . Rearranging terms for (3.6) yields

$$\begin{aligned} \frac{\alpha}{1+\lambda} &= 1 - \frac{H\left(\theta_2^{1*}\right) - \bar{b}^1}{H'\left(\theta_2^{1*} - e\left(\theta_2^{1*}; \varpi_1\right)\right)} \frac{f\left(\theta_2^{1*}\right)}{F\left(\theta_2^{1*}\right) - F\left(\theta_1^{1*}\right)} \\ &= 1 - \frac{\underline{c}_{CC}^1 - \bar{b}^1}{H'\left(H^{-1}\left(\bar{c}_2^1\right)\right)} \frac{f(\theta_2^{1*})}{\mathbb{E}(D^{CF}|W = \varpi_1)} \end{aligned} \tag{4.15}$$

with $\underline{c}_{CC}^1 \equiv H\left(\theta_2^{1*}\right)\left(> H\left(\theta_2^{1*} - \bar{e}_2^1\right) = \bar{c}_2^1\right)$ being the lower bound for period-2 cost under CC, conditioned on $W = \varpi_1$. Since all terms of (4.15) are either directly identifiable from the observed variables or already shown identifiable, (4.15) identifies $\alpha/(1+\lambda)$.

The intertemporal preference γ

γ is an additional parameter featured only in the two-period setting. Identification of γ is achieved from (3.5), which involves observations from both periods, as follows:

$$\begin{aligned}\gamma &= \frac{\bar{b}^1 - \underline{b}^1}{\bar{b}^1 - H\left(\theta_1^{1*} - \bar{e}_1^1\right) - \psi_1\left(\bar{e}_1^1\right)} \\ &= \frac{\bar{b}^1 - \underline{b}^1}{\bar{b}^1 - \bar{c}_1^1 - \psi_1\left(\bar{e}_1^1\right)} \end{aligned} \tag{4.16}$$

where $\bar{e}_1^1 \in \left[\underline{e}, \bar{e}_2^1\right]$ is identified as the solution to the f.o.c. $\psi_1'\left(\bar{e}_1^1\right) = H'\left(H^{-1}\left(\bar{c}_1^1\right)\right)$. Alternatively, according to (4.11) and the fact that $F_{C_1|W}\left(\bar{c}_1^1|\varpi_1\right) = 1$, $\bar{e}_1^1$ is identified as $\bar{e}_1^1 = \inf\left\{e : F_{E_1|W}\left(e|\varpi_1\right) = 1\right\}$.

4.5 Summary of identification results and further discussion

In the following theorem we summarize identification results established in Section 4.1 - 4.4:

Theorem 2 *Let Assumptions 3.1, 3.2, 4.1 and 4.2 hold.*
(i) In the single-period setting: $H^{-1}(\cdot)$ is nonparametrically identified on $\left[\underline{c}, \bar{c}^1\right]$. Consequently, $H(\cdot)$, $F(\cdot)$ and $(\psi_1(\cdot), \psi_2(\cdot))'$ are nonparametrically identified on $\left[\underline{\theta} - \underline{e},\ \theta^{1*} - \bar{e}^1\right]$, $\left[\underline{\theta}, \theta^{1*}\right]$ and $\left[\underline{e}, \bar{e}^1\right]$, respectively. In addition, $\alpha/(1+\lambda)$ is identified;
(ii) In the two-period setting: $H^{-1}(\cdot)$ is nonparametrically identified on $\left[\underline{c}, \bar{c}_2^1\right]$. Consequently, $H(\cdot)$, $F(\cdot)$ and $(\psi_1(\cdot), \psi_2(\cdot))'$ are nonparametrically identified on $\left[\underline{\theta} - \underline{e},\ \theta_2^{1*} - \bar{e}_2^1\right]$, $\left[\underline{\theta}, \theta_2^{1*}\right]$ and $\left[\underline{e}, \bar{e}_2^1\right]$, respectively. In addition, both $\alpha/(1+\lambda)$ and γ are identified.

Without imposing further restrictions, it is impossible to identify $H^{-1}(\cdot)$ on $\left(\bar{c}^1, \bar{c}\right]$ or equivalently $H(\cdot)$ on $\left(\theta^{1*} - \bar{e}^1, \overline{\theta}\right]$, or to identify $F(\cdot)$ on $\left(\theta^{1*}, \overline{\theta}\right]$. This is because $\left(\theta^{1*}, \overline{\theta}\right]$ is the region for θ where the ϖ_1-class agent would choose CR rather than FP, $\left(\bar{c}^1, \bar{c}\right]$ and $\left(\theta^{1*} - \bar{e}^1, \overline{\theta}\right]$ being the corresponding regions for the cost and $\theta - e^*$, respectively. And there is a lack of (structural) information under CR regions. After all, under CR, we only observe the cost, and the effort is constantly zero (rather than being strict monotonic under FP). In particular, the identification strategy for $H(\cdot)$, which is essentialized into Equations (4.2) - (4.5), only works on the region where both ϖ_1- and ϖ_2-classes choose FP, thus is not applicable to the region $\left(\theta^{1*} - \bar{e}^1, \overline{\theta}\right]$.

Supposing $H^{-1}(\cdot)$ were identified on $\left(\bar{c}^1, \bar{c}\right]$, or, equivalently, supposing $H(\cdot)$ were identified on $\left(\theta^{1*} - \bar{e}^1, \overline{\theta}\right]$, it would be straightforward to identify $F(\cdot)$ on $\left(\theta^{1*}, \overline{\theta}\right]$ via the relationship $\theta = H^{-1}(c)$ for $c \in \left(\bar{c}^1, \bar{c}\right]$ according to (4.6). In what follows, we discuss parametric identification for $H^{-1}(\cdot)$ on its entire domain $[\underline{c}, \bar{c}]$: We assume $H^{-1}(\cdot)$ to admit a parametric form $H^{-1}(\cdot;\beta)$ on $[\underline{c}, \bar{c}]$ for some finite dimensional parameter $\beta \in \mathbb{R}^{k_\beta}$. Note that this parameterization guarantees an unique extrapolation of $H^{-1}(\cdot)$ from the information-rich region $\left[\underline{c}, \bar{c}^1\right]$ to $\left(\bar{c}^1, \bar{c}\right]$. And β is identified from the parametric version of (4.4)

$$H^{-1}(c;\beta) = \underline{\theta} - \underline{e} + \sum\nolimits_{k=0}^{\infty} \Delta\tilde{E}\left(\tau_k(c) \cdot p_1\right), \text{ for all } c \in \left[\underline{\theta} - \underline{e}, \theta^{1*} - \bar{e}^1\right].$$

Alternatively, β can be identified from

$$H^{-1}\left(C\left(a_l;\varpi_1\right);\beta\right) \quad = \quad H^{-1}\left(C\left(a_l\cdot\frac{p_1}{p_2};\varpi_2\right);\beta\right)+\Delta\tilde{E}\left(a_l\cdot p_1\right), \tag{4.17}$$

for a preselected sequence of constants $\{a_1,a_2,...,a_L\}\subset(0,1)$ with large enough L, together with the condition $H^{-1}\left(\underline{c};\beta\right)=\underline{\theta}-\underline{e}$. Typically, $L\geq k_\beta$ is required for, but dose not necessarily guarantee, identification of β. Detailed conditions for identification of nonlinear parametric models (i.e., uniqueness of solution to a system of nonlinear equations) is case dependent and generally complicated. Nevertheless, for a special case in which $H(\cdot)$ is specified to take a quadratic form, which we adopt in our empirical study, identification of β is straightforward. Note that, once β is identified, we can recover $\overline{\theta}$, the upper bound for θ, as $\overline{\theta}=H^{-1}\left(\bar{c}_R;\beta\right)$ with $\bar{c}_R$ being the upper bound of the realized costs under CR.

We end this section with a few remarks regarding the parameterization of $H\left(\cdot\right)$: (i) Parameterization of various cost functions is a widely adopted in structural analysis in the literature. E.g., Luo et al. (2018) parameterize a cost function to identify the truncated distribution of consumer type; (ii) It is possible to develop consistent procedures to test a parametric specification on $H\left(\cdot\right)$. Heuristically, this could be done by comparing identification results under a given parametric form on $\left[\underline{\theta},\theta_2^{1*}\right]$ with nonparametric results on the same (sub)interval.

5 Estimation

In this section, we propose estimation procedures for the model primitives in the single-period setting, regarding which we note the followings: (i) The proposed procedures closely follow the identification strategies in Section 4.1 - 4.3; (ii) They are straightforwardly extendable to the two-period setting, which we discuss at the end of this section; (iii) As previously mentioned, we may observe other covariates Z, on which some of the model primitives potentially depend. With Z being suppressed, the estimation procedures to be presented in this section are implicitly conditioned on $Z=z$ for a given z. These procedures provide a basis for developing more sophisticated estimation procedures that incorporate Z nonparametrically, e.g. some kernel based ones. Nevertheless, depending on the sample size, or/and to avoid "curse of dimensionality", it may be preferable to incorporate Z parametrically.

The estimation is based on $\left\{c_i,q_i,w_i,x_i,y_i,d_i^F\right\}_{i=1}^n$, a random sample collected from n observed single-period contracts. And we denote by $n_f\equiv\sum_{i=1}^n d_i^F$ the number of FP contracts, and $n_f^j\equiv\sum_{i=1}^n d_i^F\cdot\mathbf{1}\left(w_i=\varpi_j\right)$ the number of FP contracts with $w_i=\varpi_j$, for $j=1,2$.

Effort distribution

Denote by $f_{E|W}\left(\cdot|\varpi_j\right)$, $f_{V_1|W}\left(\cdot|\varpi_j\right)$ and $f_{V_2|W}\left(\cdot|\varpi_j\right)$ the conditional pdf's of$E$, V_1 and V_2 on $W=\varpi_j$, respectively. We estimate $f_{E|W}\left(\cdot|\varpi_j\right)$ jointly with $f_{V_1|W}\left(\cdot|\varpi_j\right)$, $f_{V_2|W}\left(\cdot|\varpi_j\right)$ and $m\left(\cdot\right)$ by the sieve maximum likelihood estimation (SMLE) proposed in Shen (1997). Specifically, the

SMLE estimator $\left(\widehat{m}\left(\cdot\right),\widehat{f}_{E|W}\left(\cdot|\varpi_{j}\right),\widehat{f}_{V_{1}|W}\left(\cdot|\varpi_{j}\right),\widehat{f}_{V_{2}|W}\left(\cdot|\varpi_{j}\right)\right)$ are the solution to constrained optimization problem

$$\max \sum_{i=1}^{n_f^j} \ln \int f_{V_1|W}\left(x_{i2}-e|\varpi_j\right) f_{V_2|W}\left(y_{i2}-m(e)|\varpi_j\right) f_{E|W}\left(e|\varpi_j\right) de,$$

over $\left\{\left(m\left(\cdot\right),f_{E|W}\left(\cdot|\varpi_{j}\right),f_{V_{1}|W}\left(\cdot|\varpi_{j}\right),f_{V_{2}|W}\left(\cdot|\varpi_{j}\right)\right)\right\}$, subject to the following constraints: (i) $F_{E|W}\left(\cdot|\varpi_j\right)$, $f_{V_1|W}\left(\cdot|\varpi_j\right)$ and $f_{V_2|W}\left(\cdot|\varpi_j\right)$ are nonnegative, and integrate to one over corresponding supports (, hence are legit pdf's); (ii) $\int v f_{V_1|W}\left(v|\varpi_j\right) dv = \int v f_{V_2|W}\left(v|\varpi_j\right) dv = 0$ (to satisfy Assumption 4.1). Depending on the sample size, one may adopt either fully nonparametric specifications on $f_{E|W}(\cdot|\varpi_j)$, $f_{V_1|W}(\cdot|\varpi_j)$ and $f_{V_2|W}(\cdot|\varpi_j)$ or parametric/semiparametric ones on some of them. Note that, for nonparametric specifications, all the constraints can be easily implemented in appropriate linear sieve spaces, such as those spanned by Hermite orthogonal basis functions. For comprehensive studies on the implementation and properties of sieve estimations, see Shen (1997), Ai and Chen (2003), Chen (2007), Carroll et al. (2010), and Chen et al. (2014), among others.

The cost function

Here we focus on parametric estimation of $H\left(\cdot\right)$, by following the parametric identification result in Section 4.5. Specifically, we adopt a parameterization $H\left(\cdot\right) = H(\cdot;\beta)$, and estimate the parameter $\beta \in \mathbb{R}^{k_\beta}$. Motivated by (4.17), we estimated β by

$$\hat{\beta} = \underset{\beta: H(\underline{\theta}-\underline{e};\beta)=\underline{c}}{\operatorname{argmin}} \sum_{l=1}^{L} \left[H^{-1}\left(\widehat{C}\left(a_l \cdot \frac{\hat{p}_1}{\hat{p}_2};\varpi_2\right);\beta\right) - H^{-1}\left(\widehat{C}\left(a_l;\varpi_1\right);\beta\right) + \Delta\widehat{\tilde{E}}(a_l\cdot\hat{p}_1)\right]^2 \quad (5.1)$$

with $\Delta\widehat{\tilde{E}}(a_l\cdot\hat{p}_1) = \widehat{E}\left(a_l\cdot\frac{\hat{p}_1}{\hat{p}_2};\varpi_2\right) - \widehat{E}\left(a_l;\varpi_1\right)$, where $\{a_l\}_{l=1}^{L} \subset (0,1)$ is a sequence of grid points; $\widehat{C}(\tau;\varpi_j)$ and $\widehat{E}(\tau;\varpi_j)$ are estimators of $C(\tau;\varpi_j)$ and $E(\tau;\varpi_j)$, respectively, and $\hat{p}_j = \sum_{i=1}^{n} d_i^F \mathbf{1}\left(w_i=\varpi_j\right)/\sum_{i=1}^{n}\mathbf{1}\left(w_i=\varpi_j\right)$ as in (3.12)

Recall that $C(\tau;\varpi_j)$ and $E(\tau;\varpi_j)$ are defined as the τ's quantile of C conditional on $W=\varpi_j$ and that of E conditional on $W=\varpi_j$, respectively. Since both C and W are observed from the data, the estimator $\widehat{C}(\tau;\varpi_j)$ can be calculated from the corresponding empirical distribution. On the other hand, the estimator $\widehat{E}(\tau;\varpi_j)$ can be calculated based on $\widehat{f}_E\left(\cdot|\varpi_j\right)$ obtained from the SMLE. $H(\cdot)$ and $H^{-1}(\cdot)$ are then estimated as $\widehat{H}(\cdot) = H\left(\cdot;\hat{\beta}\right)$ and $\widehat{H}^{-1}(\cdot) = H^{-1}\left(\cdot;\hat{\beta}\right)$, respectively.

We note that the non/semi-parametric estimation of $H(\cdot)$ on $[\underline{\theta}-\underline{e},\theta^{1*}-\bar{e}^1]$ is possible, by following the identification strategy in Section 4.2, and is left for future study.

Type distribution and other primitives

Following (4.6), given an observed cost c_i from the sample, we estimate the corresponding type θ_i by

$$\hat{\theta}_i = \begin{cases} \widehat{H}^{-1}(c_i) + \sum_{j=1,2} \mathbf{1}(w_i = \varpi_j) \cdot \widehat{F}^{-1}_{E|W}\left(\widehat{F}_{C|W}\left(c_i|\varpi_j\right)|\varpi_j\right), \text{if } d_i^F = 1; \\ \widehat{H}^{-1}(c_i), \text{otherwise.} \end{cases}$$

Then the CDF for θ can estimated based on the sample of fitted values $\{\hat{\theta}_i\}_{i=1}^n$ as $\widehat{F}(\theta) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}(\hat{\theta}_i \leq \theta)$. And the pdf $f(\cdot)$ can be estimated by a kernel estimator based on $\{\hat{\theta}_i\}_{i=1}^n$.

Following (4.9), $\psi_j(\cdot)$ is estimated by

$$\widehat{\psi}_j(e) = b^{j*} - \bar{c}^j - \int_e^{\hat{\bar{e}}^j} H'\left(H^{-1}\left(\widehat{F}^{-1}_{C|W}\left(\widehat{F}_{E|W}\left(v|\varpi_j\right)|\varpi_j\right); \hat{\beta}\right); \hat{\beta}\right) dv. \tag{5.2}$$

Following 4.10), the ratio $\alpha/(1+\lambda)$ is estimated by

$$\widehat{\alpha/(1+\lambda)} = 1 - \frac{\underline{c}^1_{CR} - b^{1*}}{H'\left(H^{-1}\left(\bar{c}^1; \hat{\beta}\right); \hat{\beta}\right)} \frac{\widehat{f}\left(\hat{\theta}^{1*}\right)}{\hat{p}_1}, \tag{5.3}$$

where $\hat{\theta}^{1*}$ and $\hat{p}_1 \equiv \sum_{i=1}^n d_i^F \mathbf{1}(w_i = \varpi_1)/\sum_{i=1}^n \mathbf{1}(w_i = \varpi_1)$ (as defined in (3.12)) are estimators of θ^{1*} and $\mathbb{E}\left(D^F|W = \varpi_1\right)$, respectively. Specifically, since $F(\theta^{1*}) = \mathbb{E}\left(D^F|W = \varpi_1\right)$, $\hat{\theta}^{1*}$ can be constructed as $\hat{\theta}^{1*} = \widehat{F}^{-1}(\hat{p}_1)$.

Proposition 5.1 establishes regularity conditions for consistency of the main estimators.

Proposition 5.1 *Let Assumptions Assumptions 3.1, 3.2, 4.1 and 4.2 hold. In addition, let the following conditions hold: (i) $H(c;\beta)$ is continuous in both $c \in (\underline{\theta} - \underline{e}, \overline{\theta})$ and $\beta \in \mathcal{B}$ with $\mathcal{B}$ being compact; (ii) β is identified via (4.17); and (iii) $\left\{\hat{p}_1, \hat{p}_2, \widehat{C}(a_l; \varpi_1), \widehat{C}\left(a_l \cdot \frac{p_1}{p_2}; \varpi_2\right), \widehat{E}(a_l; \varpi_1), \widehat{E}\left(a_l \cdot \frac{p_1}{p_2}; \varpi_2\right)\right\}$ for $l = 1, ..., L$ are consistent. Then $\hat{\beta}$, $\widehat{F}(\cdot)$, $\widehat{\alpha/(1+\lambda)}$, and $\widehat{\psi}(\cdot)$ are consistent.*

Estimation in the two-period setting

In the two-period setting, first we need to estimate the distribution of the second-period effort, namely $F_{E_2|W}$ and $f_{E_2|W}$. The procedure for estimating $F_{E_2|W}$ and $f_{E_2|W}$ remains the same, except for that it is based on the second-period observations. The procedure for estimating the cost function remains the same as well, except for the followings: (i) $\hat{p}_j$ in (5.1) is replaced by $\hat{p}_{2,j} = \sum_{i=1}^n \left(d_i^{FF} + d_i^{CF}\right) \mathbf{1}(w_i = \varpi_j)/\sum_{i=1}^n \mathbf{1}(w_i = \varpi_j)$ as defined in (3.14), for $j = 1, 2$; (ii) The cost quantiles are estimated according the observed second-period costs, and the effort quantiles are obtained based on $\widehat{F}_{E_2|W}$. The estimation of the other model primitives follows closely to the identification results in Section 4.4, and is as follows:

Following (4.13), given an observed second-period cost c_{i2}, we estimate θ_i by

$$\hat{\theta}_i = \begin{cases} \widehat{H}^{-1}(c_{i2}) + \sum_{j=1,2} \mathbf{1}(w_i = \varpi_j) \cdot \widehat{F}^{-1}_{E_2|W} \left(\widehat{F}_{C_2|W} \left(c_{i2}|\varpi_j \right) |\varpi_j \right), \text{if } d_i^{FF} + d_i^{CF} = 1; \\ \widehat{H}^{-1}(c_{i2}), \text{otherwise.} \end{cases}$$

Then the distribution of θ can estimated based on the sample of fitted values $\{\hat{\theta}_i\}_{i=1}^n$.

Following (4.14), $\psi_j(\cdot)$ is estimated by

$$\widehat{\psi}_j(e) = \bar{b}^j - \bar{c}_2^j - \int_e^{\hat{\bar{e}}_2^j} H' \left(H^{-1} \left(\widehat{F}^{-1}_{C_2|W} \left(\widehat{F}_{E_2|W} \left(v|\varpi_j \right) |\varpi_j \right); \hat{\beta} \right); \hat{\beta} \right) dv.$$

Following (4.15), $\alpha/(1+\lambda)$ is estimated by

$$\widehat{\alpha/(1+\lambda)} = 1 - \frac{\underline{c}^1_{CC} - \bar{b}^1}{H' \left(H^{-1} \left(\bar{c}_2^1; \hat{\beta} \right); \hat{\beta} \right)} \frac{\widehat{f}\left(\hat{\theta}_2^{1*} \right)}{\tilde{p}},$$

where $\hat{\theta}_2^{1*} \equiv \widehat{F}^{-1}\left(\hat{p}_{2,1}\right)$, with $\hat{p}_{2,1} = \sum_{i=1}^n \left(d_i^{FF} + d_i^{CF} \right) \mathbf{1}\left(w_i = \varpi_1 \right) / \sum_{i=1}^n \mathbf{1}\left(w_i = \varpi_1 \right)$ (as defined in (3.14)), is an estimator of θ_2^{1*} motivated by the fact that $F(\theta_2^{1*}) = \mathbb{E}\left(D^{FF} + D^{CF} | W = \varpi_1 \right)$.

$$\tilde{p} \equiv \frac{\sum_{i=1}^n \mathbf{1}\left(w_i = \varpi_1 \right) \cdot d_i^{CF}}{\sum_{i=1}^n \mathbf{1}\left(w_i = \varpi_1 \right)}$$

serves an estimator of $\mathbb{E}\left(D^{CF} | W = \varpi_1 \right)$.

Finally, based on (4.16), the intertemporal preference γ can be estimated by

$$\hat{\gamma} = \frac{\bar{b}^1 - \underline{b}^1}{\bar{b}^1 - \bar{c}_1^1 - \widehat{\psi}_1\left(\hat{\bar{e}}_1^1 \right)},$$

where $\hat{\bar{e}}_1^1$ is an estimator of $\bar{e}_1^1$. Based on the fact that $\bar{e}_1^1 = \min\left\{ e : F_{E_1|W}\left(e|\varpi_1 \right) = 1 \right\}$ (discussed in Section 4.4), $\hat{\bar{e}}_1^1$ can be constructed as $\hat{\bar{e}}_1^1 = \widehat{F}^{-1}_{E_1|W}\left(1 - \varepsilon_n | \varpi_1 \right)$ with tuning parameter $\varepsilon_n > 0$ such that $\varepsilon_n \downarrow 0$ as $n \to \infty$.

6 Empirical illustrations

In this section, we apply our methods to analyze contracts for urban transportation services in France. The main objectives are to test the widely assumed LCC, and to evaluate how this specification assumption affects welfare assessment.

6.1 Background and the data

The urban transportation industry in France is regulated. Since 1982, for each urban area with significant size and a public transportation network, a local authority (a city, a group of cities, or a district) must operate the network directly or contract with a single operator to provide passenger transportation service. In 2010, 91% of the urban public transportation contracts are implemented by operators, and only in few large cities such as Paris and Marseille, the transportation service is fully integrated within the city administration (Heddebaut, 2017). Most of the network are operated by four major firms: Agir, Connex, Keolis, and Transdev, among which the first three are private, and the last one is semipublic.

In a French transportation contract, the local authority sets the bus route, fare structure, capacity, quality of service, etc., and the operator delivers the passenger transportation services specified in the contract. Motivated by private information of operators and limited auditing capacity of local authorities, we model the bilateral interaction between a local authority and an operator as a principal-agent problem, and assume away competition between agents.[10] In this industry, operators often have better information than local authorities about drivers' skills and behavior, fuel consumption, number of buses required for a certain route, and other cost-related factors. It also has been well documented that local authorities have difficulties to verify effort that operators put into cost-reducing and efficient management (e.g., see Domenach (1987)). Thus, the French urban transportation contracts fit our framework.

The dataset includes 543 two-period contracts implemented in France from 1987 to 2001, collected from the Centre d'Etude et de Recherche du Transport Urbain (CERTU, Lyon) surveys. For each contract in the dataset, the type (FF/CF/CC), realized cost, subsidy in any FP period are recorded. Besides, additional characteristics of operators/contracts, such as the labor fee, number of employees, number of drivers, size of rolling stock (measured by the number of vehicles), and ownership of operators, are recorded. Among these contracts, 281 are FF, 88 are CF, and the remaining 174 are CC. The same dataset is used in Gagnepain et al. (2013).

Table 1 provides summary statistics of the dataset. The average cost and subsidy are around 17 million and 19 million euros, respectively, which suggests operating the transportation network to be profitable on average to the operators. The average labor fee is 10.7 million, ac-

[10] Before 1993, the usual practice for local authorities was to simply award contracts with one operator via negotiation, and renewed the contracts after five years. In 1993, the 'Sapin' Act was promulgated, in hope of preventing corruption affairs and enhancing competition among operators. The Act mandates a three-step procedure for a local authority to select an operator: (i) The local authority selects qualified operators; (ii) The selected operators submit bids; (iii) The authority chooses one of the operators to negotiate with, and eventually sign a FP-CR contract with it. However, local authorities are neither obliged to choose the best (lowest) bid as in a standard procurement auction, nor required to state selection criteria in their documents. As a consequence, competition among operators plays minimal role in the French urban transportation industry, which is evidenced by the fact that, in more than 60% of auctions, only one operator submitted its bid, and for 90% of contracts the local authority renewed the contract with the incumbent instead of switching to another operator (Gautier and Yvrande-Billon, 2013). Therefore, following the existing literature (e.g., Gagnepain and Ivaldi (2002)), we assume away such competition, and apply the current principal-agent framework to the empirical study. Nevertheless, in our current framework, the parameter α, which measures the bargaining power of the operator, potentially captures competition, if there is any, among potential operators (Gagnepain et al., 2013). And we believe that the simplified framework, together with the parameter α, provides a reasonable first-order approximation to the industry, although it is an *ad hoc* approach and may lose some generality.

Table 1: Summary Statistics

Variables	Mean	Std. Dev.	Median	Min	Max
NO. of Contracts	543				
NO. of FF	281				
NO. of CF	88				
NO. of CC	174				
Cost	16860	15954	10347	2397	93993
Subsidy	18794	18236	12039	2265	114483
NO. of employees	413	364	267	68	1772
NO. of drivers	278	216	144	47	1182
Labor fee	10740	10241	6650	716	53178
Rolling stock	165	121	84	33	724
Private owenship	0.52	0.50	1.00	0.00	1.00

Note: The units of cost, subsidy and labor fee are in 1000 euros.

counting for 64 percent of the total cost. This suggests that reducing the labor fee is critical to increase the operator's profit. On average, 278 out of 413 (so more than a half) employees are drivers, suggesting the transportation industry to be labor-intensive. As to the ownership, over half of the contracts are signed with privately-owned operators.

6.2 Testing for the LCC in the French transportation industry

First, we examine whether the LCC holds (i.e., whether operators' cost function is linear) for the French transportation industry. Specifically, we test the testable implication (3.10) of the LCC, by applying the inference procedure proposed in Section 3.2. Recall that the procedure requires a binary variable W, which is independent of firms' innate cost θ but potentially affects the disutility of effort. Our choice of W is the ownership dummy, which equals one if an operator is privately owned, and equals zero otherwise. Note that the innate cost θ captures factors that affect a firm's (in)efficiency but are inherently not under its control. In the transportation industry, these factors are mainly technologies related to urban-transportation that are accessible to an operator firm. And it is reasonable to assume that accessibility to this kind of technologies has little to do with ownership status. On the other hand, a firm can counterbalance a high innate cost by executing managerial efforts, such as monitoring drivers, providing training programs (to promote driving habits that increase fuel efficiency), and solving potential conflicts, etc. The empirical evidence from Gagnepain et al. (2013) shows that Transdev's effort is less costly than the other three firms, suggesting that the disutility function is affected by the ownership and that public-owned firms incur less disutility.[11] Thus we will continue to use the ownership

[11]Among the four operators in the dataset, only Transdev is semipublic, while all other three are private.

dummy as the exclusion variable W in Section 6.3 for further empirical analysis.

Based on all 369 FF and CF contracts, the test rejects the null hypothesis of a constant optimal effort at a significance level of 1%. Since, as we discussed, the constant optimal effort is a direct implication of the LCC, this testing result suggests that the LCC fails to adequately describe the cost structure in the French transportation industry. Instead, one should adopt a nonlinear specification for the cost function in analyzing the French transportation contracts, as what we do next in the empirical study.

6.3 The welfare analysis

In the welfare analysis, our objective is to evaluate how the LCC, as a specification assumption, affects the assessment of welfare. Specifically, we estimate welfare of the contracts with and without imposing the LCC, then conduct a comparison.

The assessment of welfare plays a central role in empirical studies of contracts. An example is to evaluate welfare loss associated with renegotiation relative to full-commitment. In the current study, we maintain that renegotiation is allowed, which is indeed the case in the urban transportation industry in France, and evaluate how econometric specifications of cost function, i.e., LCC and non-LCC, affect welfare analysis of the contracts. Our empirical finding in the welfare analysis provides helpful insights on resolving the issue of evaluating the welfare losses and related ones: Our comparison between the welfare assessments with and without the LCC shows a noticeable difference. This finding suggests that the LCC, if misspecified, would cause a substantial bias in the assessment of welfare assessment. Therefore, one should be very cautious when deciding whether or not to impose the LCC, especially when there is empirical evidence against it, such as the rejection of its testable implication(s) in a hypothesis test.

As a pre-step for the welfare assessments, we estimate the cost function, distribution of innate cost and other model primitives with (possibly) nonlinear cost in Section 6.3.1, and we estimate the same primitives under the LCC in Section 6.3.2. Based on the estimation results, we conduct welfare analysis in Section 6.3.3.

6.3.1 Estimation with (possibly) nonlinear cost

In the estimation, we choose to use the share of drivers among all employees and the labor fee as X and Y, i.e., the two proxies of effort, respectively. More specifically, since employees mainly consist of drivers and engineers, X is constructed as

$$X = \frac{\text{NO. of drivers}}{\text{NO. of drivers} + \text{ NO. of engineers}} = 1 - \text{ share of engineers.}$$

These choices for X and Y are plausible because: (i) The share of engineers provides a measure of endowed skills of an operator, thus is expected to be negatively related to the innate cost,[12]

[12] As Gagnepain et al. (2013) point out, engineers are generally responsible for research and development, quality control, maintenance, and efficiency.

which in turn is positively related to the optimal effort. Therefore, we expect X (i.e., $1-$ share of engineers) to be positively related to the optimal effort; (ii) Our choice for Y is in accordance with Cicala (2015)'s suggestion that one can employ cost-related variables to infer agent's effort. Specifically, the labor fee, which represents 64 percent of the total cost, can be reasonably interpreted as a function of the optimal effort due to the one-to-one mapping between the optimal effort and the innate cost. We still use the ownership dummy as the exclusion variable W. And we take into account the operator's rolling stock (i.e., NO. of transit vehicles) as covariate Z.

Recall that our identification strategies only identify the operator's bargaining power α and the cost of public funds λ up to the ratio $\alpha/(1+\lambda)$ (as in (2.16)). Nevertheless, for ease of discussion, we set $\lambda = 0.3$ as in Gagnepain et al. (2013), which follows the empirical suggestion by Ballard et al. (1985) that $\lambda \in [0.15, 0.40]$ in an efficient tax systems. This arrangement enables further identification and estimation of α.

Semiparametric specifications

Given the moderate sample size, we adopt a semiparametric specification for the estimation: (i) All probability distributions are treated fully nonparametrically; (ii) We parameterize the cost and disutility functions, $m(\cdot)$ of Assumption 4.1, as well as the dependence of θ on Z. Specifically, we set

$$\begin{aligned} H(\theta - e) &= \beta_1(\theta - e) + \beta_2(\theta - e)^2, && (6.1) \\ \psi_j(e) &= \kappa_1^j e + \kappa_2^j e^2, && (6.2) \\ m_j(e) &= \zeta_1^j e + \zeta_2^j e^2, && (6.3) \end{aligned}$$

where the sub/sup-script j indicates 'being conditioned on $W = \varpi_j$', for $j = 1, 2$. We set

$$\theta = \theta_0 + \lambda_\theta Z, \quad (6.4)$$

where θ_0 follows an unknown distribution on $[0, 1]$, and λ_θ is an unknown parameter. The parameterizations above introduce the following parameters: β_1, β_2, λ_θ, and κ_1^j, κ_2^j, ζ_1^j, ζ_2^j, for $j = 1, 2$.

The parametric specifications of (6.1), (6.2) and (6.4) can be shown to imply the optimal effort (under FP) to take the form

$$E = E_0^j + \lambda_e^j Z, \text{ conditioned on } W = \varpi_j, \quad (6.5)$$

where E_0^j is a random variable independent of Z, and λ_e is a constant.[13] The distribution of E_0^j and λ_e are not jointly identifiable, thus we normalize the support of E_0^j to $[-1, 1]$. Equation (6.5)

[13]It can be shown that e_0^j is determined by θ_0, β_1, β_2, κ_1^j and κ_2^j, and that λ_e^j is determined by λ_θ, β_1, β_2, κ_1^j and κ_2^j. Consequently, $E_0^j = e_0(\theta_0; \beta_1, \beta_2, \kappa_1^j, \kappa_2^j)$, and $\lambda_e^j = \lambda_e(\beta_1, \beta_2, \kappa_1^j, \kappa_2^j, \lambda_\theta)$, for $j = 1, 2$.

further implies the following to hold for $f_{E|W,Z}$, the conditional distribution of E on (W, Z):

$$f_{E|W,Z}\left(e|\varpi_j, z\right) = f_{E_0|W}\left(e - \lambda_e^j z|\varpi_j\right), \tag{6.6}$$

which we utilize (6.6) in the estimation of $f_{E|W,Z}$.

The estimation closely follows the procedures proposed in Section 5, with some minor adjustments that adapt to the parameterization of (6.1) - (6.4). We provide details on the estimation in the Appendix B.

Estimation results

We collect main estimation results in Table 2 . As shown in the table, the intertemporal weight γ is estimated at 0.03, which is much smaller 0.97 (i.e., $1 - 0.03$). This suggests that operators pay relatively much more attention to the second-period profit. The operator's bargaining power is estimated at 1.281, which is consistent with the theoretical restriction that $\alpha < 1 + \lambda$. As to the parameters introduced by parameterization (6.1) - (6.4) (i.e., β_1, β_2, λ_θ, and κ_1^j, κ_2^j, ζ_1^j, ζ_2^j for $j = 1, 2$), the estimation results for these parameters are to be used in the welfare assessment, and should not be over-interpreted per se. Though it is worth noting that a one-sided test on β_2 ($\mathrm{H}_0 : \beta_2 = 0$ vs $\mathrm{H}_1 : \beta_1 > 0$) rejects the null hypothesis at a significance level of 2.5%. This suggests the cost function to be convex, which is consistent with our theoretic setting.

6.3.2 Estimation with linear cost

According to the LCC, we parameterize the cost function as $H(\theta - e) = \beta(\theta - e)$. Such a linear specification makes part of the estimation simpler. As discussed in Section 3.2, a linear cost function induces a constant optimal effort under FP, regardless of θ. Thus, the two effort proxies X and Y (for identifying/estimating the distribution of effort) are no longer needed, as the distribution now degenerates to a single point (with probability mass one) for a give W. However the lack of variation in optimal effort makes it impossible to identify the disutility functions $\{\psi_1(\cdot), \psi_2(\cdot)\}$ without further assumptions. For this reason, we normalize $\psi_1(e) = e^2$, and adopt a simple parameterization $\psi_2(e) = \kappa e^2$ for ψ_2, which satisfies Assumption 3.2, and is widely used in the related literature, such as Laffont and Tirole (1988), Rogerson (2003), Chu and Sappington (2007), and Battaglini (2007).[14] The estimation results are reported in Table 3. Like those in Table 2, these results are to be used in the welfare assessment.

6.3.3 Welfare assessments and comparison

The objective here, as noted earlier, is to evaluate how the specification of cost function affects the welfare assessment. For the evaluation, we focus on the average social welfare SW.

[14] As shown in Appendix C (the last paragraph), under a slightly more general specification $\psi_2(e) = \kappa_1 e + \kappa_2 e^2$, $\left(e^{2*}, \kappa_1, \kappa_2\right)$ are not jointly identifiable.

Table 2: Estimation results with nonlinear cost

Functions	Parameters	Private($W = \varpi_1$)	Public ($W = \varpi_2$)
Cost	Ineff. (β_1)	-9.003	
		(60.621)	
	Ineff. $\times$ Ineff. (β_2)	31425.6**	
		(14884.8)	
Disutility	Effort (κ_1)	39063.0***	44276.6***
		(12799.2)	(13828.7)
	Effort $\times$ Effort (κ_2)	2003.2	1830.0
		(2544.1)	(7346.1)
Measurement Y	Effort (ζ_1)	11.151***	7.789
		(4.274)	(5.901)
	Effort $\times$ Effort (ζ_2)	10.788***	17.991***
		(4.230)	(6.692)
	Rolling stock (λ_θ)	-0.0004	
		(0.0007)	
	Intertemporal weight (γ)	0.0296	
		(0.1024)	
	Bargaining power (α)	1.2805***	
		(0.3103)	

Standard errors in parentheses are bootstrapped 1000 times;
*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$;
λ_θ, r and α are first estimated conditioned on $W = \varpi_1$ and $W = \varpi_2$ separately. What is reported here are the averages.

Table 3: Estimation results with linear cost

Parameters	Cost β	Disutility κ	Intertemporal weight γ	Bargaining power α
Est.	24761.7	41863.9	0.2148***	1.2999
	(20509.1)	(156297.1)	(0.0942)	(6.8898)

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

The assessment of SW is delicate in the sense that SW consists of quite a few components, each of which needs to be estimated separately and then put together. Nevertheless, the assessment is straightforward since it follows the general definition (2.5) tidily as we show below. Besides, most of SW's components have already been estimated in Section 6.3.1 and 6.3.2.

Denote by $SW(z)$ the welfare conditional on $Z=z$. And denote by $SW^j(z)$ the welfare conditional on $Z=z$ and $W=\varpi_j$, for $j=1,2$. By (2.5), we have

$$\begin{aligned} SW^j(z) &= S-(1+\lambda)Q^j(z)+\alpha U^j(z), \\ SW(z) &= \sum_{j=1,2} SW^j(z)\cdot \Pr(W=\varpi_j) \equiv S-(1+\lambda)Q(z)+\alpha U(z), \end{aligned}$$

with $Q(z)\equiv\sum_{j=1,2}Q^j(z)\cdot\Pr(W=\varpi_j)$ and $U(z)\equiv\sum_{j=1,2}U^j(z)\cdot\Pr(W=\varpi_j)$, where

$$\begin{aligned} Q^j(z) &\equiv \underline{b}^j F_{\theta|z}\left(\theta_1^{j*}(z)|z\right)+\int_{\theta_1^{j*}(z)}^{\theta_2^{j*}(z)}\left[rH(\theta)+(1-r)\bar{b}^j\right]dF_{\theta|z}(\theta|z) \\ &\quad+\int_{\theta_2^{j*}(z)}^{1+\lambda_\theta z}H(\theta)\,dF_{\theta|z}(\theta|z), \end{aligned} \tag{6.7}$$

$$\begin{aligned} U^j(z) &\equiv \int_{\lambda_\theta z}^{\theta_1^{j*}(z)}\left[\underline{b}^j-H(\theta-e(\theta))-\psi(e(\theta))\right]dF_{\theta|z}(\theta|z) \\ &\quad+(1-r)\int_{\theta_1^{j*}(z)}^{\theta_2^{j*}(z)}\left[\bar{b}^j-H(\theta-e(\theta))-\psi(e(\theta))\right]dF_{\theta|z}(\theta|z), \end{aligned} \tag{6.8}$$

for $j=1,2$, according to (2.12). $\theta_1^{j*}(z)$ and $\theta_2^{j*}(z)$ represent the two cutoff values for the innate cost conditioned on $Z=z$ and $W=\varpi$, and they need to be estimated in order to obtain $\widehat{Q}^j(z)$ and $\widehat{U}^j(z)$, i.e., estimators of $Q^j(z)$ and $U^j(z)$. We provide estimation details for $\theta_1^{j*}(z)$ and $\theta_2^{j*}(z)$ in Appendix B. Finally,

$$SW = \int_z SW(z)\,dF_Z(z) \equiv S-(1+\lambda)Q+\alpha U$$

with

$$\begin{aligned} Q &\equiv \int_z Q(z)\,dF_Z(z) = \int_z\sum_{j=1,2}Q^j(z)\cdot\Pr(W=\varpi_j)\,dF_Z(z), \\ U &\equiv \int_z U(z)\,dF_Z(z) = \int_z\sum_{j=1,2}U^j(z)\cdot\Pr(W=\varpi_j)\,dF_Z(z). \end{aligned}$$

Accordingly, we assess SW, up to a constant S, by

$$\widehat{SW}=S-\left(1+\hat{\lambda}\right)\int_{\mathcal{Z}}\widehat{Q}(z)\,d\widehat{F}_Z(z)+\hat{\alpha}\int_{\mathcal{Z}}\widehat{U}(z)\,d\widehat{F}_Z(z),$$

with $\widehat{Q}(z)=\sum_{j=1,2}\widehat{Q}^j(z)\cdot\widehat{\Pr}(W=\varpi_j)$ and $\widehat{U}(z)=\sum_{j=1,2}\widehat{U}^j(z)\cdot\widehat{\Pr}(W=\varpi_j)$, where $\widehat{\Pr}(W=\varpi_j)=\sum_{i=1}^n\mathbf{1}(w_i=\varpi_1)/n$. And $\widehat{Q}^j(z)$ and $\widehat{U}^j(z)$ are constructed by plugging in corresponding estimators to (6.7) and (6.8), respectively. Under the LCC, SW is assessed by plugging in the estimators from Section 6.3.2. Without the LCC, SW is assessed by plugging in the estimators from Section 6.3.1, which are obtained under a specification that allows for a nonlinear cost structure.

Table 4: Assessments of average welfare under different specifications

	Model specification	Assessment
$(1+\lambda)\,Q$	Nonlinear cost	2.79
(i.e., social cost)	Linear cost	4.61
	Difference ($\Delta\,(1+\lambda)\,Q$)	1.82
αU	Nonlinear cost	2.72
(i.e., informational rent)	Linear cost	2.95
	Difference ($\Delta\alpha U$)	0.23
Welfare difference	$\Delta SW = \Delta\alpha U - \Delta\,(1+\lambda)\,Q$	-1.59

All estimates are in million euros.

We report the welfare assessments and comparison results in Table 4. As indicated, whether or not the LCC is imposed makes a substantial difference in assessing the average social welfare. For the French transportation industry, the difference is estimated at a magnitude of 1.59 million euros, which is almost 10% of the average contract cost (16.86 million euros by Table 1). This finding suggests the welfare assessment to be sensitive to the specification of cost function. Thus, one needs to be cautious on deciding whether or not to impose linearity, since mis-specification could lead to substantial bias. In addition, we find the difference in estimated social cost to contribute the major portion of the difference in welfare assessment. The average social is estimated much higher under linear cost than that under nonlinear cost (4.61 versus 2.79, the former being 65% more). We note that these findings are reminiscent of An and Zhang (2018), who show in theory that, for single-period contracts, the social welfare can differ a lot depending on whether the cost function is linear or not.

7 Conclusion

We provide a rigorous econometric framework for analyzing FP-CR contracts in both the single-period and two-period settings. We prove that the model is nonparametrically identified on intervals corresponding to FF and CF contracts. Further we provide 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 propose a semiparametric procedure to estimate the model primitives. In the empirical study, based on data from public transportation procurement contracts in France, we find strong evidence against linearity of the cost function. The importance of this empirical finding is further evidenced by a welfare analysis, which shows the welfare assessment to be sensitive to the specification of cost function.

Our identification strategy can be potentially extended to handle linear cost sharing (LCS) contracts, which encompass FP and CR as special cases. Intuitively, observing the more flexible payment from LCS contracts (compared with that of FP or CR contracts) would provide additional information that is rich enough to identify the additional finite dimensional parameter of LCS contracts. We leave this extension for future research.

APPENDIX

The appendix is organized as follows: In Appendix A, we provide proofs of the main results. In Appendix B, we provide details on the estimation involved in the empirical study. In Appendix C, we provide proofs of Lemmas A.1 - A.3, which are employed (in Appendix A) for proving the main results, and provide additional supplementary material.

A Proofs of the main results

In the main context of the paper, the main results are stated, in order of occurrence, as Lemma 2.1, Propositions 2.1, 2.2, Corollary 3.1, Theorem 1, Lemma 4.1, Propositions 4.1, 4.2, Theorem 2 and Proposition 5.1.

We present proofs of these results in the same order as listed above. Note that Theorem 2 simply summarizes identification results stated in Proposition 4.2 and Section 4.3-4.4, all of which are proven/explained elsewhere. Thus, a dedicated proof for Theorem 2 is unnecessary.

In proving the main results, we utilize the following technical lemmas, labeled as Lemmas A.1 – A.3, whose proofs are provided in Appendix C.

Lemma A.1 (Renegotiation-proof) *There is no loss of generality in restricting the analysis to contracts of the form* $\mathbf{b} = (b_1^{FF}, b_2^{FF}, b_2^{CF})$ *that come unchanged through the renegotiation process, that is,* $R \equiv (b_2^{FF}, b_2^{CF})$ *maximizes the principal's second-period welfare subject to the following acceptance conditions:*

$$\tilde{b}_2^{FF} \geq b_2^{FF} \quad \text{and} \quad \tilde{b}_2^{CF} \geq b_2^{CF}. \tag{A.1}$$

Lemma A.2 *An initial offer* $\mathbf{b} = \left(b_1^{FF}, b_2^{FF}, b_2^{CF}\right)$ *is renegotiation-proof if and only if the following two conditions hold:*

$$\theta_2\left(b_2^{CF}\right) \geq \theta_1\left(\mathbf{b}\right); \tag{A.2}$$

$$\left(1 - \frac{\alpha}{1+\lambda}\right) \frac{F\left(\theta_2\left(b_2^{CF}\right)\right) - F\left(\theta_1\left(\mathbf{b}\right)\right)}{f\left(\theta_2\left(b_2^{CF}\right)\right)} \quad \geq \quad \frac{\left[H\left(\theta_2\left(b_2^{CF}\right)\right) - b_2^{CF}\right]}{H'\left(\theta_2\left(b_2^{CF}\right) - e\left(\theta_2\left(b_2^{CF}\right)\right)\right)}, \tag{A.3}$$

where $\theta_1\left(\mathbf{b}\right)$ *is the cutoff type between FF and CF induced by* C*, and* $\theta_2\left(b_2^{CF}\right)$ *is the cutoff one between CF and CC induced by* C*. Note that, as its notation suggests,* $\theta_2\left(b_2^{CF}\right)$ *depends on* b_2^{CF}*, but not* b_1^{FF} *or* b_2^{FF}*, under renegotiation-proof.*

Lemma A.3 (Uniform weak convergence) *Let Assumption 3.1, 3.2, and Condition T.1 hold.*
(i) It holds that

$$\begin{aligned}
&\sqrt{n_f}\left\{\left[\widehat{C}\left(\tau;\varpi_1\right) - \widehat{C}\left(\tau p;\varpi_2\right)\right] - \left[\widehat{C}\left(\nu;\varpi_1\right) - \widehat{C}\left(\nu p;\varpi_2\right)\right]\right\} \\
&- \sqrt{n_f}\left\{\left[C\left(\tau;\varpi_1\right) - C\left(\tau p;\varpi_2\right)\right] - \left[C\left(\nu;\varpi_1\right) - C\left(\nu p;\varpi_2\right)\right]\right\} \\
\xrightarrow{\mathcal{L}}\ &(1,0)\, z\left(\tau\right)/J\left(\tau\right) - (1,1)\, z\left(\tau p\right)/J\left(\tau p\right) - (1,0)\, z\left(\nu\right)/J\left(\nu\right) + (1,1)\, z\left(\nu p\right)/J\left(\nu p\right)
\end{aligned}$$

where: (i) $J\left(\tau\right) \equiv \mathbb{E}\left[f_{C|W}\left(q_1\left(\tau\right) + q_2\left(\tau\right)W|W\right)W\right]$ *for any given* $\tau \in [0,1]$ *with*

$$\left(q_1\left(\tau\right), q_2\left(\tau\right)\right) \quad \equiv \quad \underset{q_1,q_2}{\operatorname{argmin}}\, \mathbb{E}\left[\rho_\tau\left(c_{f_i} - q_1 - q_2 \cdot w_{f_i}\right)\right];$$

(ii) $z\left(\cdot\right)$ *is a zero mean tight Gaussian process on* $L^\infty[0,1]$*; (iii) The weak convergence "*$\xrightarrow{\mathcal{L}}$*" is in the sense of Chapter 1.3 in van der Vaart and Wellner (1996).*

Proof of Lemma 2.1. As discussed in the main context, since the CR option provides no incentive to exert effort, the optimal effort is always zero under CR.

Taking first-order derivative w.r.t. θ on both side of the f.o.c. (2.2) (i.e., $H'(\theta - e\left(\theta\right)) = \psi'(e\left(\theta\right)))$ yields

$$e'\left(\theta\right) = \frac{H''\left(\theta - e\left(\theta\right)\right)}{H''\left(\theta - e\left(\theta\right)\right) + \psi''\left(e\left(\theta\right)\right)} \in \left(0,1\right), \tag{A.4}$$

where the strict inequalities follow from the conditions $H''\left(\cdot\right) > 0$ and $\psi''\left(\cdot\right) > 0$ by Assumption 2.1. (A.4) verifies the second claim.

It follows from (A.4) and $H''\left(\cdot\right) > 0$ that

$$\begin{aligned}
\frac{d\left[\theta - e\left(\theta\right)\right]}{d\theta} \quad &= \quad 1 - e'\left(\theta\right) \ > \ 0, \\
\frac{dH\left(\theta - e\left(\theta\right)\right)}{d\theta} \quad &= \quad H''\left(\theta - e\left(\theta\right)\right)\left[1 - e'\left(\theta\right)\right] \ > \ 0,
\end{aligned}$$

which verifies the third claim. ■

Proof of Proposition 2.1. For any given fixed price b within suitable range,[15] there exists a unique cut-off value $\theta(b)$ s.t. an agent with $\theta = \theta(b)$ is indifferent between FP (with fixed price b) and CR, i.e., these two options lead to the same amount of payoff for him. Recall from (2.3) that the payoff under FP is

$$u_F(\theta, b) \;=\; b - H(\theta - e(\theta)) - \psi(e(\theta)).$$

Also recall that, under the normalizing condition $\psi(0) = 0$ by Assumption 2.1, the payoff under CR would always be zero. Therefore, we have

$$u_F(\theta(b), b) \;=\; b - H(\theta(b) - e(\theta(b))) - \psi(e(\theta(b))) \;=\; 0. \tag{A.5}$$

As a special case of (A.5) above, we have

$$u_F(\theta^*, b^*) \;=\; b^* - H(\theta^* - e(\theta^*)) - \psi(e(\theta^*)) \;=\; 0$$

with $\theta^* \equiv \theta(b^*)$, which verifies (2.8) in Proposition 2.1 (ii).

To verify Proposition 2.1 (i), note that, for a given b, we have

$$\frac{\partial u_F(\theta, b)}{\partial \theta} \;=\; -\frac{dH(\theta - e(\theta))}{d\theta} - \psi'(e(\theta)) \cdot \frac{de(\theta)}{d\theta} \;<\; 0,$$

where the inequality follows from that $\frac{dH(\theta - e(\theta))}{d\theta} > 0$ and $\frac{de(\theta)}{d\theta} > 0$ (according to Lemma 2.1) and that $\psi'(\cdot) > 0$ (by Assumption 2.1 (ii)). Consequently, $u_F(\theta, b) > u_F(\theta(b), b) = 0$ for all $\theta < \theta(b)$, and $u_F(\theta, b) < u_F(\theta(b), b) = 0$ for all $\theta > \theta(b)$. In particular, for $b = b^*$, we have $u_F(\theta, b^*) > u_F(\theta^*, b^*) = 0$ for all $\theta < \theta^*$, and $u_F(\theta, b^*) < u_F(\theta^*, b^*) = 0$ for all $\theta > \theta^*$, which verify Proposition 2.1 (i).

We verify (2.9) in Proposition 2.1 (ii) as follows: According to (A.5), we can rewrite $\tilde{\pi}(\theta, b)$ (specified by (2.6)) as

$$\tilde{\pi}(\theta, b) \;=\; \begin{cases} S - (1+\lambda) b + \alpha [b - H(\theta - e(\theta)) - \psi(e(\theta))], \text{ for } \theta \le \theta(b); \\ S - (1+\lambda) H(\theta), \text{ for } \theta > \theta(b). \end{cases}$$

In turn, we can rewrite $\pi(b)$ (specified by (2.7)) as

$$\begin{aligned} \pi(b) \;&=\; \int_{\underline{\theta}}^{\bar{\theta}} \tilde{\pi}(\theta, b)\, dF(\theta) \\ &=\; S - (1+\lambda) F(\theta(b))\, b + \alpha F(\theta(b))\, b + \alpha \int_{\underline{\theta}}^{\theta(b)} [H(\theta - e(\theta)) + \psi(e(\theta))]\, dF(\theta) \\ &\quad - (1+\lambda) \int_{\theta(b)}^{\bar{\theta}} H(\theta)\, dF(\theta). \end{aligned}$$

[15] Specifically, $b \in \left[H(\underline{\theta} - e(\underline{\theta})) + \psi(e(\underline{\theta})), H(\bar{\theta} - e(\bar{\theta})) + \psi(e(\bar{\theta}))\right]$. Note that, if b is below this range, CR would be preferred by all θ. Similarly, if b is above this range, FP would be preferred by all θ.

Since $b^* = \underset{b}{\operatorname{argmax}}\, \pi(b)$, the following f.o.c. holds:

$$\begin{aligned} \left.\frac{d\pi(b)}{db}\right|_{b=b^*} &= [\alpha - (1+\lambda)]\, F(\theta^*) - (1+\lambda)\, f(\theta^*)\, \theta^{'}(b^*)\, b^* + (1+\lambda)\, H(\theta^*)\, f(\theta^*)\, \theta^{'}(b^*) \\ &= 0, \end{aligned}$$

for which rearranging terms yields

$$\left(1 - \frac{\alpha}{1+\lambda}\right) \frac{F(\theta^*)}{f(\theta^*)} = [H(\theta^*) - b^*]\, \theta^{'}(b^*)\,. \tag{A.6}$$

Note that taking derivative w.r.t. b on both side of (A.5) and rearranging terms yield

$$\begin{aligned} \theta^{'}(b) &= \left\{H^{'}(\theta(b) - e(\theta(b))) - \left[H^{'}(\theta(b) - e(\theta(b))) - \psi^{'}(e(\theta(b)))\right] e^{'}(\theta(b))\right\}^{-1} \\ &= \frac{1}{H'(\theta(b) - e(\theta(b)))}, \end{aligned}$$

where the second equality is due to that $H^{'}(\theta(b) - e(\theta(b))) - \psi^{'}(e(\theta(b))) = 0$ according to (2.2). As a special case of the equation above, for $b = b^*$, we have

$$\theta^{'}(b^*) = \frac{1}{H'(\theta^* - e(\theta^*))}. \tag{A.7}$$

Substituting (A.7) into (A.6) yields

$$\left(1 - \frac{\alpha}{1+\lambda}\right) \frac{F(\theta^*)}{f(\theta^*)} = \frac{H(\theta^*) - b^*}{H'(\theta^* - e(\theta^*))},$$

which completes the proof. ∎

Proof of Proposition 2.2.

According to (2.10) – (2.12), the expected intertemporal welfare when offering $\mathbf{b} = \left(b_1^{FF}, b_2^{FF}, b_2^{CF}\right)$ is

$$\begin{aligned} \pi_{int}(\mathbf{b}) = & \int_{\underline{\theta}}^{\theta_1(\mathbf{b})} \left\{S - (1+\lambda)(\gamma b_1^{FF} + (1-\gamma) b_2^{FF})\right. \\ & \left. + \alpha\left[\gamma b_1^{FF} + (1-\gamma) b_2^{FF} - H(\theta - e(\theta)) - \psi(e(\theta))\right]\right\} dF(\theta) \\ & + \int_{\theta_1(\mathbf{b})}^{\theta_2\left(b_2^{CF}\right)} \left\{S - (1+\lambda)(\gamma H(\theta) + (1-\gamma) b_2^{CF})\right. \\ & \left. + \alpha(1-\gamma)\left[b_2^{CF} - H(\theta - e(\theta)) - \psi(e(\theta))\right]\right\} dF(\theta) \\ & + \int_{\theta_2\left(b_2^{CF}\right)}^{\bar{\theta}} [S - (1+\lambda) H(\theta)]\, dF(\theta). \end{aligned} \tag{A.8}$$

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

$$\max_{\mathbf{b}} \pi_{int}(\mathbf{b}) \quad \text{subject to (A.3)}. \tag{A.9}$$

[16]We assume (A.2) holds with strict inequality and (A.3) holds with equality as shown in Lemma A.2.

According to Lemma A.2, the optimal fixed prices $\mathbf{b}^{\dagger}=\left(b_1^{FF\dagger},b_2^{FF\dagger},b_2^{CF\dagger}\right)$ and the corresponding cut-off types $(\theta_1\left(\mathbf{b}^{\dagger}\right),\theta_2\left(b_2^{CF\dagger}\right))$ satisfy

$$b_1^{FF\dagger}+\frac{(1-\gamma)}{\gamma}\left(b_2^{FF\dagger}-b_2^{CF\dagger}\right)=H\left(\theta_1\left(\mathbf{b}^{\dagger}\right)-e\left(\theta_1\left(\mathbf{b}^{\dagger}\right)\right)\right)+\psi\left(e\left(\theta_1\left(\mathbf{b}^{\dagger}\right)\right)\right), \tag{A.10}$$

$$b_2^{CF\dagger}=H\left(\theta_2\left(b_2^{CF\dagger}\right)-e\left(\theta_2\left(b_2^{CF\dagger}\right)\right)\right)+\psi\left(e\left(\theta_2\left(b_2^{CF\dagger}\right)\right)\right). \tag{A.11}$$

Due to the fact that $\partial\theta_1\left(\mathbf{b}^{\dagger}\right)/\partial b_1^{FF}=[\gamma/(1-\gamma)]\,\partial\theta_1\left(\mathbf{b}^{\dagger}\right)\partial b_2^{FF}$ by (A.10), it can be shown that the f.o.c.'s for $b_1^{FF\dagger}$ and $b_2^{FF\dagger}$ are identical, and hence lead to the same optimal solution $b_1^{FF\dagger}=b_2^{FF\dagger}\equiv\underline{b}$.

Next, define $\bar{b}\equiv b_2^{CF\dagger}$. Following similar arguments as in the discussion of Proposition 2 in Gagnepain et al. (2013), it can be shown that the renegotiation-proof constraint (A.3) is binding in the equilibrium (i.e., at $\mathbf{b}^{\dagger}$). Consequently, the f.o.c.'s with respect to $b_1^{FF\dagger}$ and $b_2^{CF\dagger}$ leads to

$$\frac{\gamma\left(1+\lambda-\alpha\right)F\left(\theta_1\left(\mathbf{b}^{\dagger}\right)\right)}{f\left(\theta_1\left(\mathbf{b}^{\dagger}\right)\right)} = \frac{(1+\lambda)\left[\gamma H\left(\theta_1\left(\mathbf{b}^{\dagger}\right)\right)+(1-\gamma)\,\bar{b}-\underline{b}\right]-\vartheta\left(1-\frac{\alpha}{1+\lambda}\right)}{H'\left(\theta_1\left(\mathbf{b}^{\dagger}\right)-e\left(\theta_1\left(\mathbf{b}^{\dagger}\right)\right)\right)},$$

$$\begin{aligned}\left(1-\frac{\alpha}{1+\lambda}\right)\left[F\left(\theta_2\left(b_2^{CF\dagger}\right)\right)-F\left(\theta_1\left(\mathbf{b}^{\dagger}\right)\right)\right] =\;& \frac{\left[H\left(\theta_2\left(b_2^{CF\dagger}\right)\right)-\bar{b}\right]f\left(\theta_2\left(b_2^{CF\dagger}\right)\right)}{H'\left(\theta_2\left(b_2^{CF\dagger}\right)-e\left(\theta_2\left(b_2^{CF\dagger}\right)\right)\right)}\\ &-\frac{\left[\gamma H\left(\theta_1\left(\mathbf{b}^{\dagger}\right)\right)+\bar{b}(1-\gamma)-\underline{b}\right]f\left(\theta_1\left(\mathbf{b}^{\dagger}\right)\right)}{\gamma H'\left(\theta_1\left(\mathbf{b}^{\dagger}\right)-e\left(\theta_1\left(\mathbf{b}^{\dagger}\right)\right)\right)}\\ &+\frac{\vartheta\cdot m\left(\theta_1\left(\mathbf{b}^{\dagger}\right),\theta_2\left(b_2^{CF\dagger}\right),\lambda,\gamma,\alpha\right)}{(1+\lambda)(1-\gamma)},\end{aligned}$$

where

$$\begin{aligned}&m\left(\theta_1\left(\mathbf{b}^{\dagger}\right),\theta_2\left(b_2^{CF\dagger}\right),\lambda,\gamma,\alpha\right)\\ &\equiv\left(1-\frac{\alpha}{1+\lambda}\right)\left(\frac{f\left(\theta_2\left(b_2^{CF\dagger}\right)\right)}{H'\left(\theta_2\left(b_2^{CF\dagger}\right)-e\left(\theta_2\left(b_2^{CF\dagger}\right)\right)\right)}-\frac{f\left(\theta_1\left(\mathbf{b}^{\dagger}\right)\right)(\gamma-1)}{\gamma H'\left(\theta_1\left(\mathbf{b}^{\dagger}\right)-e\left(\theta_1\left(\mathbf{b}^{\dagger}\right)\right)\right)}\right)\\ &-\frac{\left[H'\left(\theta_2\left(b_2^{CF\dagger}\right)\right)f\left(\theta_2\left(b_2^{CF\dagger}\right)\right)+\left(H\left(\theta_2\left(b_2^{CF\dagger}\right)\right)-\bar{b}\right)f'\left(\theta_2\left(b_2^{CF\dagger}\right)\right)\right]H'\left[\theta_2\left(b_2^{CF\dagger}\right)-e\left(\theta_2\left(b_2^{CF\dagger}\right)\right)\right]}{\left[H'\left(\theta_2\left(b_2^{CF\dagger}\right)-e\left(\theta_2\left(b_2^{CF\dagger}\right)\right)\right)\right]^3}\\ &+\frac{H''\left[\theta_2\left(b_2^{CF\dagger}\right)-e\left(\theta_2\left(b_2^{CF\dagger}\right)\right)\right]\left[1-e^{*\prime}\left(\theta_2\left(b_2^{CF\dagger}\right)\right)\right]\left[H\left(\theta_2\left(b_2^{CF\dagger}\right)\right)-\bar{b}\right]f\left(\theta_2\left(b_2^{CF\dagger}\right)\right)}{\left[H'\left(\theta_2\left(b_2^{CF\dagger}\right)-e\left(\theta_2\left(b_2^{CF\dagger}\right)\right)\right)\right]^3},\end{aligned}$$

and $\vartheta>0$ is the Lagrange multiplier corresponding to the binding constraint (A.3), which completes the proof of Proposition 2.2. ■

Proof of Corollary 3.1. Corollary 3.1 simply restates the results in Propositions 2.1 and 2.2, but conditioned on W. And its proof follows the same logic as the proofs of these two propositions, so is skipped. ■

Proof of Theorem 1. If $\mathbb{H}_0$ is true, then $C(\tau;\varpi_1) - C(\tau p;\varpi_2) \equiv$ a constant for $\tau \in [0,1]$ almost sure. Consequently, $C(\tau;\varpi_1) - C(\tau p;\varpi_2) = C(\nu;\varpi_1) - C(\nu p;\varpi_2)$ almost sure. And it immediately follows from Lemma A.3 that, under $\mathbb{H}_0$

$$\begin{aligned} & \sqrt{n_f}\left\{\left[\widehat{C}(\tau;\varpi_1) - \widehat{C}(\tau p;\varpi_2)\right] - \left[\widehat{C}(\nu;\varpi_1) - \widehat{C}(\nu p;\varpi_2)\right]\right\} \\ = & \sqrt{n_f}\left\{\left[\widehat{C}(\tau;\varpi_1) - \widehat{C}(\tau p;\varpi_2)\right] - \left[\widehat{C}(\nu;\varpi_1) - \widehat{C}(\nu p;\varpi_2)\right]\right\} \\ & - \sqrt{n_f}\left\{[C(\tau;\varpi_1) - C(\tau p;\varpi_2)] - [C(\nu;\varpi_1) - C(\nu p;\varpi_2)]\right\} \\ \xrightarrow{\mathcal{L}} & (1,0)\, z(\tau)/J(\tau) - (1,1)\, z(\tau p)/J(\tau p) - (1,0)\, z(\nu)/J(\nu) + (1,1)\, z(\nu p)/J(\nu p). \end{aligned} \tag{A.12}$$

According to Theorem 1.11.1 (Extended continuous mapping) in van der Vaart and Wellner (1996), it follows from (A.12) that, under $\mathbb{H}_0$

$$\begin{aligned} S_n \equiv & \int_0^1 n_f \left\{\left[\widehat{C}(\tau;\varpi_1) - \widehat{C}(\tau p;\varpi_2)\right] - \left[\widehat{C}(\nu;\varpi_1) - \widehat{C}(\nu p;\varpi_2)\right]\right\}^2 d\tau \\ \xrightarrow{\mathcal{L}} & \int_0^1 \left[(1,0)\, z(\tau)/J(\tau) - (1,1)\, z(\tau p)/J(\tau p) - (1,0)\, z(\nu)/J(\nu) + (1,1)\, z(\nu p)/J(\nu p)\right]^2 d\tau. \end{aligned}$$

Note that the asymptotic distribution above is unaffected when p is replaced by a $\sqrt{n}$-consistent estimator $\hat{p}$, which verifies Part (i) of Theorem 1.

To verify Part (ii) of Theorem 1, define $\Delta(\tau) \equiv [C(\tau;\varpi_1) - C(\tau p;\varpi_2)] - [C(\nu;\varpi_1) - C(\nu p;\varpi_2)]$. It follows from the support for $C|W$ being bounded that $\Delta(\cdot)$ is bounded on $[0,1]$. Moreover, under any fixed alternative, $|\Delta(\cdot)|$ is bounded away from zero on a set with nonzero measure. Consequently, we have

$$0 < \int_0^1 [\Delta(\tau)]^2\, d\tau < +\infty. \tag{A.13}$$

It follows from (A.13), the weak law of large number and the stochastic equicontinuity implied by Lemma A.3 that

$$\left[\widehat{C}(\tau;\varpi_1) - \widehat{C}(\tau p;\varpi_2)\right] - \left[\widehat{C}(\nu;\varpi_1) - \widehat{C}(\nu p;\varpi_2)\right] \xrightarrow{p} \Delta(\tau)$$

uniformly in $\tau \in [0,1]$, which, according to Theorem 1.11.1 in van der Vaart and Wellner (1996), implies that

$$\frac{1}{n_f} S_n \xrightarrow{p} \int_0^1 [\Delta(\tau)]^2\, d\tau. \tag{A.14}$$

Note that the convergence in probability result above is unaffected when p in S_n is replaced by a $\sqrt{n}$-consistent estimator $\hat{p}$, in which case S_n is changed to T_n. I.e., $\frac{1}{n_f}T_n \xrightarrow{p} \int_0^1 [\Delta(\tau)]^2 d\tau$, which, together with (A.13), show that T_n diverges to ∞ in probability at a rate of $O_p(n_f)$. This completes the proof. ■

Proof of Lemma 4.1. Denote by $(\widetilde{C}, \widetilde{D}^{FF}, \widetilde{D}^{CF}, \widetilde{\overline{B}}, \widetilde{\underline{B}})$ the underlying variables that correspond to $\widetilde{\mathcal{S}}$. 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 $(D^{FF}, D^{CF}, C, \overline{B}, \underline{B}) = (\widetilde{D}^{FF}, \widetilde{D}^{CF}, \widetilde{C}, \widetilde{\overline{B}}, \widetilde{\underline{B}})$, and that the equality (2.16) 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 $\widetilde{\theta}$, $\widetilde{\theta} \leq \widetilde{\theta}_l$ is equivalent to $\theta \leq \theta_l$, which implies that $\widetilde{D}^{FF} = D^{FF}$. Similarly, we have $\widetilde{D}^{CF} = D^{CF}$. Note that

$$\widetilde{\psi}'(\widetilde{e}^*) = \widetilde{H}'(\widetilde{\theta} - \widetilde{e}^*) \Rightarrow \psi'[(\widetilde{e}^* - \xi_0)/\xi_1] = H'[(\widetilde{\theta} - \widetilde{e}^*)/\xi_1], \tag{A.15}$$

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

$$\begin{aligned}
\widetilde{C} &= \widetilde{H}(\widetilde{\theta} - \widetilde{e}(\widetilde{\theta}^*)) = H[(\xi_1\theta - \xi_1 e(\theta^*))/\xi_1] = H(\theta - e(\theta^*)) = C, && \text{(A.16)}\\
\widetilde{\overline{B}} &= \widetilde{H}(\widetilde{\theta}_u - \widetilde{e}(\widetilde{\theta}_u)) + \widetilde{\psi}(\widetilde{e}(\widetilde{\theta}_u)) = H((\xi_1\theta_u - \xi_1 e(\theta_u))/\xi_1) + \psi(\xi_1 e(\theta^*)/\xi_1) = \overline{B}, \\
\widetilde{\underline{B}} &= r[\widetilde{H}(\widetilde{\theta}_l - \widetilde{e}(\widetilde{\theta}_l)) + \widetilde{\psi}(\widetilde{e}(\widetilde{\theta}_l))] + (1-r)\overline{B} \\
&= r[H(\theta_l - e(\theta_l)) + \psi(e(\theta_l))] + (1-r)\overline{B} = \underline{B}. && \text{(A.17)}
\end{aligned}$$

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

$$\left(1 - \frac{\alpha}{1+\lambda}\right) \frac{\widetilde{F}(\widetilde{\theta}_u) - \widetilde{F}(\widetilde{\theta}_l)}{\widetilde{f}(\widetilde{\theta}_u)} = \xi_1 \left(1 - \frac{\alpha}{1+\lambda}\right) \frac{F(\theta_u) - F(\theta_l)}{f(\theta_u)} = \xi_1 \frac{H(\theta_u) - \overline{B}}{H'(\theta_u - e(\theta_u, w))} \tag{A.18}$$

and

$$\frac{\widetilde{H}(\widetilde{\theta}_u) - \widetilde{\overline{B}}}{\widetilde{H}'(\widetilde{\theta}_u - \widetilde{e}(\widetilde{\theta}_u))} = \frac{\widetilde{H}(\widetilde{\theta}_u) - \widetilde{\overline{B}}}{\widetilde{\psi}'(\widetilde{e}(\widetilde{\theta}))} = \frac{H(\theta_u) - \overline{B}}{\psi'(e^*)/\xi_1} = \xi_1 \frac{H(\theta_u) - \overline{B}}{H'(\theta_u - e(\theta_u))}. \tag{A.19}$$

Hence,

$$\left(1 - \frac{\alpha}{1+\lambda}\right) \frac{\widetilde{F}(\widetilde{\theta}_u) - \widetilde{F}(\widetilde{\theta}_l)}{\widetilde{f}(\widetilde{\theta}_u)} = \frac{\widetilde{H}(\widetilde{\theta}_u) - \widetilde{\overline{B}}}{\widetilde{H}'(\widetilde{\theta}_u - \widetilde{e}(\widetilde{\theta}_u))}. \tag{A.20}$$

This completes the proof. ■

Proof of Proposition 4.1. 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 4.1, we need to impose a few regularity conditions to achieve identification, as follows. Note that these conditions are not restrictive for our setting.
(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 4.1 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 4.2. According to Corollary 3.1, under FP, the range of $C|W = \varpi_j$, i.e., realized costs conditioned on $W = \varpi_j$, is as follows:

$$\begin{aligned} C|W = \varpi_1 \in [C(0;\varpi_1), C(1;\varpi_1)] &= \left[H(\underline{\theta} - e(\underline{\theta};\varpi_1)), H\left(\theta^{1*} - e\left(\theta^{1*};\varpi_1\right)\right)\right], \\ C|W = \varpi_2 \in [C(0;\varpi_2), C(1;\varpi_2)] &= \left[H(\underline{\theta} - e(\underline{\theta};\varpi_2)), H\left(\theta^{2*} - e\left(\theta^{2*};\varpi_2\right)\right)\right]. \end{aligned}$$

Note that it follows from Assumption 4.2 (ii) and the f.o.c. (2.2) (i.e., $H'(\theta - e(\theta)) = \psi'(e(\theta))$) that $e(\underline{\theta};\varpi_1) = e(\underline{\theta};\varpi_2)$. Consequently, it holds that $C(0;\varpi_1) = C(0;\varpi_2) = \underline{C}$.

Next, we verify following claim: For any given $\tau \in (0, p_1]$ it holds that,

$$e(\theta(\tau);\varpi_1) < e(\theta(\tau);\varpi_2) \quad \text{or, equivalently,} \quad E(\tau/p_1;\varpi_1) < E(\tau/p_2;\varpi); \tag{A.21}$$

$$c(\theta(\tau);\varpi_1) > c(\theta(\tau);\varpi_2) \quad \text{or, equivalently,} \quad C(\tau/p_1;\varpi_1) > C(\tau/p_2;\varpi_2). \tag{A.22}$$

(A.21) can be verified by contradiction, as follows: Suppose $e(\theta(\tau);\varpi_1) \geq e(\theta(\tau);\varpi_2)$, which immediately implies that $\theta(\tau) - e(\theta(\tau);\varpi_1) \leq \theta(\tau) - e(\theta(\tau);\varpi_2)$. It then follows from $H''(\cdot) > 0$ (by Assumption 3.2 (i)) that

$$H'(\theta(\tau) - e(\theta(\tau);\varpi_1)) \leq H'(\theta(\tau) - e(\theta(\tau);\varpi_2)),$$

which, together with (2.2) (i.e., $H'(\theta - e(\theta)) = \psi'(e(\theta))$), imply that

$$\psi_1'(e(\theta(\tau);\varpi_1)) \leq \psi_2'(e(\theta(\tau);\varpi_2)) < \psi_1'(e(\theta(\tau);\varpi_2)), \tag{A.23}$$

where the strict inequality follows from Assumption 4.2 (ii) and the fact that $e(\theta(\tau);\varpi_2) > e(\underline{\theta};\varpi_2) = \underline{e}$ for any $\tau > 0$. However, since $\psi_1''(\cdot) > 0$ (by Assumption 3.2 (ii)), it is impossibly for both $e(\theta(\tau);\varpi_1) \geq e(\theta(\tau);\varpi_2)$ and $\psi_1'(e(\theta(\tau);\varpi_1)) < \psi_1'(e(\theta(\tau);\varpi_2))$ (as immediately implied by (A.23)) to hold. Thus, we reach a contradiction, and it has to be the case that $e(\theta(\tau);\varpi_1) < e(\theta(\tau);\varpi_2)$, as claimed by (A.21). Given that (A.21) has just been verified, (A.22) follows immediate from (A.21) and the fact that $c(\theta(\tau);\varpi_j) = H(\theta(\tau) - e(\theta(\tau);\varpi_j))$ for all $\tau \leq p_1 (= \min_{j=1,2} p_j)$.

Now equipped with (A.21) and (A.22), we proceed with the proof. Note that, for any $c \in [C(0;\varpi_1), C(1;\varpi_1)]$, there exist a unique $\tau_0(c) \in [0,1]$ s.t. $c = C(t_0(c);\varpi_1)$. Moreover, $\tau_0(c) = 0$ for $c = \underline{C} = C(0;\varpi_1)$, and $\tau_0(c) \in (0,1]$ for $c \in (C(0;\varpi_1), C(1;\varpi_1)]$.

Consequently, for a given $c \in (C(0;\varpi_1), C(1;\varpi_1)]$, it follows from (A.21) that

$$c = C(\tau_0(c);\varpi_1) > C(\tau_0(c) \cdot p_1/p_2;\varpi_2).$$

Similarly, there exists a unique $\tau_1(c) \in (0, \tau_0(c))$ s.t.

$$C(\tau_1(c);\varpi_1) = C(\tau_0(c).p_1/p_2;\varpi_2) > C(\tau_1(c).p_1/p_2;\varpi_2).$$

Iteratively, given $\tau_k(c) \in (0, \tau_{k-1}(c))$, there exists a unique $\tau_{k+1}(c) \in (0, \tau_k(c))$ s.t.

$$C(\tau_{k+1}(c);\varpi_1) = C(\tau_k(c).p_1/p_2;\varpi_2) > C(\tau_{k+1}(c).p_1/p_2;\varpi_2). \tag{A.24}$$

And we end up obtaining a strictly decreasing sequence $\{\tau_k(c)\}_k \subseteq (0,1]$. The boundedness and monotonicity of $\{\tau_k(c)\}_k$ imply that a limit for the sequence exits as $k \to \infty$, denoted by

$$\underline{\tau}(c) \quad \equiv \quad \lim_{k\to\infty} \tau_k(c).$$

By the continuity of $C(\cdot;\varpi_j)$ for both $j=1,2$ (implied by the conditional distribution of C on W being continuous), taking limit on both side of the equality from (A.24), i.e., $C(\tau_{k+1}(c);\varpi_1) = C(\tau_k(c).p_1/p_2;\varpi_2)$, yields

$$C(\underline{\tau}(c);\varpi_1) = C(\underline{\tau}(c).p_1/p_2;\varpi_2), \tag{A.25}$$

which is true only if $\underline{\tau}(c)=0$. Therefore, we have just shwon that $\tau_k(c)\to 0$ as $k\to\infty$, for any given $c\in(C(0;\varpi_1), C(1;\varpi_1)]$.

To summarize, for any given $c\in[C(0;\varpi_1), C(1;\varpi_1)]$:
(i) If $c = C(0;\varpi_1)$, then we have

$$H^{-1}(c) \quad = \quad H^{-1}(\underline{C}) \;=\; \underline{\theta} - \underline{e}.$$

(ii) If $c\in(C(0;\varpi_1), C(1;\varpi_1)]$, then we have

$$\begin{aligned}
H^{-1}(c) &= H^{-1}(C(\tau_0(c);\varpi_1)) \\
&= H^{-1}(C(\tau_0(c)\cdot p_1/p_2;\varpi_2)) + \Delta\tilde{E}(\tau_0(c)\cdot p_1) \\
&= H^{-1}(C(\tau_1(c);\varpi_1)) + \Delta\tilde{E}(\tau_0(c)\cdot p_1) \\
&= H^{-1}(C(\tau_1(c)\cdot p_1/p_2;\varpi_2)) + \Delta\tilde{E}(\tau_1(c)\cdot p_1) + \Delta\tilde{E}(\tau_0(c)\cdot p_1) \\
&= \cdots \\
&= H^{-1}(C(\tau_m(t)\cdot p_1/p_2;\varpi_2)) + \sum_{k=0}^{m}\Delta\tilde{E}(\tau_k(c)\cdot p_1).
\end{aligned}$$

Rearranging terms for the equation above yields

$$\sum_{k=0}^{m}\Delta\tilde{E}(\tau_k(c)\cdot p_1) \quad = \quad H^{-1}(c) - H^{-1}(C(\tau_m(c)\cdot p_1/p_2;\varpi_2)). \tag{A.26}$$

Letting $m\to\infty$ on both side of (A.26) and rearranging terms yields

$$\begin{aligned}
H^{-1}(c) &= H^{-1}(C(\underline{\tau}(c)\cdot p_1/p_2;\varpi_2)) + \sum_{k=0}^{\infty}\Delta\tilde{E}(\tau_k(c)\cdot p_1) \\
&= H^{-1}(C(0;\varpi_2)) + \sum_{k=0}^{\infty}\Delta\tilde{E}(\tau_k(c)\cdot p_1) \\
&= \underline{\theta} - \underline{e} + \sum_{k=0}^{\infty}\Delta\tilde{E}(\tau_k(c)\cdot p_1),
\end{aligned}$$

where the second equation follows from $\underline{\tau}(c) = \lim_{k\to\infty} \tau_k(c) = 0$, which we have already shown. This completes the proof. ■

Proof of Proposition 5.1. The proposed $\{\hat{\theta}_i\}_{i=1}^n$, $\widehat{F}(\cdot)$, $\widehat{\alpha/(1+\lambda)}$, $\widehat{\psi}(\cdot)$, and $\hat{\delta}$ are all plug-in type estimators based on the estimated cost function $\widehat{H}(\cdot)$. So the consistency of $\widehat{H}(\cdot)$ is the key to guaranteeing consistency of all other estimators. Since, under the parameterization $H(\cdot) = H(\cdot;\beta)$, the cost function is estimated by $\widehat{H}(\cdot) = H\left(\cdot;\hat{\beta}\right)$, we focus to verify the consistency of $\hat{\beta}$ in this proof.

Define criterion functions

$$\begin{aligned} Q(\beta) &= \sum_{l=1}^{L}\left[H^{-1}\left(C(a_l\cdot\frac{p_1}{p_2};\varpi_2);\beta\right) - H^{-1}\left(C(a_l;\varpi_1);\beta\right) + E\left(a_l\cdot\frac{p_1}{p_2};\varpi_2\right) - E\left(a_l;\varpi_1\right)\right]^2, \\ Q_n(\beta) &= \sum_{l=1}^{L}\left[H^{-1}\left(\widehat{C}(a_l\cdot\frac{\hat{p}_1}{\hat{p}_2};\varpi_2);\beta\right) - H^{-1}\left(\widehat{C}(a_l;\varpi_1);\beta\right) + \widehat{E}\left(a_l\cdot\frac{\hat{p}_1}{\hat{p}_2};\varpi_2\right) - \widehat{E}\left(a_l;\varpi_1\right)\right]^2. \end{aligned}$$

The true value β_0 and the proposed estimator $\hat{\beta}$ can be equivalently characterized as

$$\beta_0 = \underset{\beta:H(\underline{\theta}-\underline{e};\beta)=\underline{c}}{\operatorname{argmin}} Q(\beta) \quad \text{and} \quad \hat{\beta} = \underset{\beta:H(\underline{\theta}-\underline{e};\beta)=\underline{c}}{\operatorname{argmin}} Q_n(\beta),$$

respectively.

Note that, under point-identification, β_0 is the unique minimizer of $Q(\cdot)$ over the parameter space $\mathcal{B}$, and $Q(\beta_0) = 0$. In addition, $Q(\cdot)$ is continuous, which is implied by the continuity of $H(\cdot)$ and the effort distribution. Consequently, for any given $\epsilon > 0$, it follows from the compactness of $\mathcal{B}$ and the continuity of $Q(\cdot)$ that

$$\inf_{\beta\in\mathcal{B}:d(\beta,\beta_0)\geq\epsilon} Q(\beta) > 0 = Q(\beta_0). \tag{A.27}$$

Also note that $\{\hat{p}_1, \hat{p}_2\}$ are estimators directly based on corresponding empirical distributions, and that $\left\{\widehat{C}(\cdot;\varpi_1), \widehat{C}(\cdot;\varpi_2), \widehat{E}(\cdot;\varpi_1), \widehat{E}(\cdot;\varpi_2)\right\}$ are obtained from standard quantile regression. Their consistency are guarantee by the random sample condition and a few regularity conditions (for the quantile regression) which are satisfied in our econometric setting. Consequently, the compactness of $\mathcal{B}$, the continuity of $H(\cdot)$, and the consistency of $\Big\{\hat{p}_1, \hat{p}_2, \widehat{C}(a_l;;\varpi_1),$ $\widehat{C}(a_l\cdot\frac{p_1}{p_2};\varpi_2), \widehat{E}(a_l;\varpi_1), \widehat{E}(a_l\cdot\frac{p_1}{p_2};\varpi_2)\Big\}_{l=1}^{L}$ guarantee that

$$\sup_{\beta\in\mathcal{B}}|Q_n(\beta) - Q(\beta)| \xrightarrow{p} 0. \tag{A.28}$$

(A.27) and (A.28) together imply that $\hat{\beta} \xrightarrow{p} \beta_0$ according to 5.7 Theorem in van der Vaart (1999). ■

B Details on estimation

The estimation mostly follows the procedures described in Section 5, and accordingly are all based on observations from period-2, except for $\alpha/(1+\lambda)$ and γ whose estimation requires observations from both periods. We still suppress the subscript t for simplicity of notations. Nevertheless, we do make some minor adjustments to Section 5's procedures, mainly in the estimation of $f_{E|W,Z}$, which adapt to the parameterization of (6.1) - (6.4). Here we provide details on the adjustments.

We estimate $f_{E|W,Z}$ according to (6.6) as the following plug-in estimator :

$$\hat{f}_{E|W,Z}\left(e|\varpi_j,z\right)=\hat{f}_{E_0|W}\left(e-\hat{\lambda}_e^j z|\varpi_j\right).$$

This is a *plug-in type estimator* based on $\hat{\lambda}_e^j$ and $\hat{f}_{E_0|W}\left(\cdot|\varpi_j\right)$: (i) $\hat{\lambda}_e^j$ is an OLS estimator of λ_e^j based on the FP subsample with $w_i=\varpi_j$, according to the linear equation $x_i=\lambda_0^j+\lambda_e^j z_i+u_i^j$ that follows directly from Assumption 4.1 and (6.5) with $u_i^j\equiv v_{1i}+e_{0i}^j-\mathbb{E}\left(E_0^j\right)$; (ii) $\hat{f}_{E_0|W}\left(\cdot|\varpi_j\right)$ is an estimator of $f_{E_0|W}\left(\cdot|\varpi_j\right)$, obtained via SMLE jointly with estimators of nuisance functions $f_{V_1|W}\left(\cdot|\varpi_j\right)$, $f_{V_2|W}\left(\cdot|\varpi_j\right)$ and $m_j\left(\cdot\right)$ (parameterized as in (6.3)). I.e., $\left\{\hat{f}_{E_0|W}\left(\cdot|\varpi_j\right),\hat{f}_{V_1|W}\left(\cdot|\varpi_j\right),\hat{f}_{V_2|W}\left(\cdot|\varpi_j\right),\hat{m}_j\left(\cdot\right)\right\}$ jointly maximize the objective function

$$\sum_{i=1}^{n_f^j}\ln\int_{-1}^{1}f_{V_1|W}\left(x_i-e_0-\hat{\lambda}_e^j z_i|\varpi_j\right)f_{V_2|W}\left(y_i-m_j(e_0+\hat{\lambda}_e^j z_i)\right)f_{E_0|W}(e_0|\varpi_j)de_0,$$

where n_f^j is the number of FP contracts with $w_i=\varpi_j$. And the density functions $f_{V_1|W}$, $f_{V_2|W}$ and $f_{E_0|W}$ are approximated by nonlinear sieves

$$f_{V_1|W}\left(\cdot|\varpi_j\right)=\left[\sum_{k=0}^{k_{1n}}a_{jk}q_k(\cdot)\right]^2,\ f_{V_2|W}\left(\cdot|\varpi_j\right)=\left[\sum_{k=0}^{k_{2n}}b_{jk}q_k(\cdot)\right]^2,\ f_{E_0|W}\left(\cdot|\varpi_j\right)=\left[\sum_{k=0}^{k_{3n}}c_{jk}q_k(\cdot)\right]^2,$$

for $j=1,2$, using *orthogonal Hermite* basis functions $q_k(\iota)=\sqrt{\frac{1}{\sqrt{\pi}k!2^k}}H_k(\iota)e^{-\frac{\iota^2}{2}}$, with $H_0(\iota)=1$, $H_1(\iota)=2\iota$ and $H_{k+1}(\iota)=2\iota H_k(\iota)-2kH_{k-1}(\iota)$. We set $k_{1n}=k_{2n}=k_{3n}=3$.

The estimation of other model primitives follows the procedures in Section 5. And we skip the details.

Estimation of $\theta_1^{j*}(z)$ and $\theta_2^{j*}(z)$

According to (6.7) and (6.8), both $Q^j\left(z\right)$ and $U^j\left(z\right)$ are dependent on $\theta_1^{j*}(z)$ and $\theta_2^{j*}(z)$. So $\theta_1^{j*}(z)$ and $\theta_2^{j*}(z)$ need to be estimated in order to obtain $\widehat{Q}^j\left(z\right)$ and $\widehat{U}^j\left(z\right)$, which are, in turn, needed for the welfare assessment. The estimation details for $\theta_1^{j*}(z)$ and $\theta_2^{j*}(z)$ are as follows.

Recall that $\theta_1^{j*}(z)$ and $\theta_2^{j*}(z)$ represent the two cutoff values for the efficiency type conditioned on $Z=z$ and $W=\varpi_j$. When the LCC is not imposed, it follows from the f.o.c. (2.2) and the parameterization (6.1) - (6.4) that

$$
\begin{aligned}
\theta_1^{j*}(z) &= \frac{\kappa_1^j - \beta_1 + 2\left(\beta_2 + \kappa_2^j\right)\bar{e}_1^j(z)}{2\beta_2} \\
\theta_2^{j*}(z) &= \frac{\kappa_1^j - \beta_1 + 2\left(\beta_2 + \kappa_2^j\right)\bar{e}_2^j(z)}{2\beta_2}
\end{aligned}
$$

with

$$
\bar{e}_1^j(z) = e\left(\theta_1^{j*}(z); \varpi_j\right) = F_{E_1|W,Z}^{-1}\left(1|\varpi_j, z\right) \quad \text{and} \quad \bar{e}_2^j(z) = e\left(\theta_2^{j*}(z); \varpi_j\right) = F_{E_2|W,Z}^{-1}\left(1|\varpi_j, z\right)
$$

where $F_{E_t|W,Z}\left(\cdot|\varpi_j, z\right)$ is the CDF of period-t effort under FP conditioned on $W = \varpi_j$ and $Z = z$, for $t = 1, 2$ and $j = 1, 2$. Accordingly, we can estimate $\theta_t^{j*}(z)$ for $t = 1, 2$ as

$$
\widehat{\theta}_t^{j*}(z) = \frac{\hat{\kappa}_1^j - \hat{\beta}_1 + 2\left(\hat{\beta}_2 + \hat{\kappa}_2^j\right)\widehat{\bar{e}}_t^j(z)}{2\hat{\beta}_2}
$$

with $\widehat{\bar{e}}_1^j(z) = \widehat{F}_{E_1|W,Z}^{-1}\left(1 - \varepsilon_n|\varpi_j, z\right)$ with tuning parameter $\varepsilon_n > 0$ such that $\varepsilon_n \downarrow 0$ as $n \to \infty$.

Similarly, under the LCC, it follows from the f.o.c. (2.2) and the parameterization specified in Section 6.3.2 that

$$
\theta_1^{j*}(z) = \bar{c}_1^j\left(z\right)/\beta + e^{j*}\left(z\right) \quad \text{and} \quad \theta_2^{j*}(z) = \bar{c}_2^j\left(z\right)/\beta + e^{j*}\left(z\right),
$$

where $e^{j*}\left(z\right)$ is the optimal effort conditioned on $W = \varpi_j$ and $Z = z$. $\bar{c}_t^j\left(z\right)$ is the cost upper bound under FP in period t, again, conditioned on $W = \varpi_j$ and $Z = z$. Note that, as its notation suggests, $e^{j*}\left(z\right)$ is invariant in θ due to the LCC, yet is possibly dependent on W and Z. It also follows from the f.o.c. (2.2) and the parameterization specified in Section 6.3.2 that

$$
\underline{\theta}(z) = \underline{c}_t^j\left(z\right)/\beta + e^{j*}\left(z\right), \text{ for } t = 1, 2.
$$

with $\underline{\theta}(z) = \lambda_\theta z$ according to (6.4) and the specification that $\theta_0 \in [0, 1]$. $\underline{c}_t^j\left(z\right)$ is the cost lower bound in period t conditioned on $W = \varpi_j$ and $Z = z$. Note that both $\bar{c}_t^j\left(z\right)$ and $\underline{c}_t^j\left(z\right)$ can be readily estimated since C, W and Z are all observed. Consequently, $e^{j*}\left(z\right)$ can be estimated as

$$
\widehat{e}^{j*}\left(z\right) = \hat{\lambda}_\theta z - \widehat{\underline{c}}_t^j\left(z\right)/\hat{\beta}.
$$

$\theta_t^{j*}(z)$ for $t = 1, 2$ can be then estimated as

$$
\widetilde{\theta}_t^{j*}(z) = \widehat{\bar{c}}_t^j\left(z\right)/\hat{\beta} + \widehat{e}^{j*}\left(z\right).
$$

References

Aaker, D. and J. Myers, *Advertising management* (Prentice-Hall, 1987).

Abito, J. M., "Measuring the welfare gains from optimal incentive regulation," *The Review of Economic Studies* 87 (2020), 2019–2048.

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

An, Y. and D. Zhang, "Simple Menus of Cost-based Contracts with Monotone Optimal Efforts," *Journal of Public Economic Theory* 20 (2018), 742–756.

Angrist, J., V. Chernozhukov and I. Fernández-Val, "Quantile Regression under Misspecification, with an Application to the U.S. Wage Structure," *Econometrica* 74 (2006), 539–563.

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

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

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.

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.

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

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

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

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.

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

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

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

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

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

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

D’Haultfœuille, X. and P. Février, “The provision of wage incentives: A structural estimation using contracts variation,” *Quantitative Economics* 11 (2020), 349–397.

Domenach, O., “Métamorphose des rapports contractuels dans les transports publics urbains: de l’autorité concédante à l’autorité organisatrice et du concessionnaire au prestataire de service,” *Transports* 32 (1987), 112–120.

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

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.

Gautier, A. and A. Yvrande-Billon, “Contract renewal as an incentive device. An application to the French urban public transport sector,” *Review of Economics and Institutions* 4 (2013), 29.

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

Guerre, E., I. Perrigne and Q. Vuong, “Nonparametric Identification of Risk Aversion in First-Price Auctions Under Exclusion Restrictions,” *Econometrica* 77 (2009), 1193–1227.

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

Heddebaut, O., *The evolution of public transport contracts in France*, Ph.D. thesis, IFSTTAR-Institut Français des Sciences et Technologies des Transports (2017).

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.

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

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

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

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

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

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

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

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

Rogerson, W. P., "An economic framework for analyzing DoD profit policy," Technical Report, RAND Corp Santa Monica, CA, 1992.

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

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.

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

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

van der Vaart, A., *Asymptotic Statistics* (Cambridge University Press, 1999).

van der Vaart, A. W. and J. Wellner, *Weak Convergence and Empirical Processes: With Applications to Statistics* (Springer, New York, 1996).

Online Supplement for
"A Structural Analysis of Simple Contracts"

Yonghong An[a], Shengjie Hong[b], Daiqiang Zhang[c]
[a] Department of Economics, Texas A&M University
[b] School of Economics and Management, Tsinghua University
[c] Department of Economics, University at Albany

This online supplement contains the proofs of technical lemmas A.1 – A.3 in listed Appendix A, an alternative identification result for $H(\cdot)$, as well as a discussion on identification under the LCC.

C Proofs of Lemmas A.1 – A.3, and supplementary material

Proof of Lemma A.1. The proof follows the same steps as the proof of Proposition 3 in Gagnepain et al. (2013), except for a more complex formula for calculating the expected welfare (detailed by (2.10) – (2.12 in Section 2.3). Such a difference in calculating the expected welfare is due to that a general cost function is in place of the special form $H(\theta - e) = \theta - e$ adopted by Gagnepain et al. (2013), yet is nonessential for proving Lemma A.1. Thus, the proof is skipped. ■

Proof of Lemma A.2. Consider any initial contract $\mathbf{b}^0 = \left(b_1^{FF,0}, b_2^{FF,0}, b_2^{CF,0}\right) \equiv \left(b_1^{FF,0}, R^0\right)$ and the renegotiated offer $\widetilde{R} = \left(\widetilde{b}_2^{FF}, \widetilde{b}_2^{CF}\right)$ that satisfies (A.1).By offering C^0 followed by a renegotiated $\widetilde{R}$, the principal anticipate the expected welfare for the second period to be

$$\begin{aligned} SW_2\left(\mathbf{b}^0, \widetilde{R}\right) &= \int_{\underline{\theta}}^{\theta_1\left(\mathbf{b}^0\right)} \left[S - (1+\lambda)\widetilde{b}_2^{FF} + \alpha\left(\widetilde{b}_2^{FF} - H\left(\theta - e(\theta)\right) - \psi\left(e(\theta)\right)\right)\right] dF(\theta) \\ &+ \int_{\theta_1\left(\mathbf{b}^0\right)}^{\theta_2\left(\widetilde{b}_2^{CF}\right)} \left[S - (1+\lambda)\widetilde{b}_2^{CF} + \alpha\left(\widetilde{b}_2^{CF} - H\left(\theta - e(\theta)\right) - \psi\left(e(\theta)\right)\right)\right] dF(\theta) \\ &+ \int_{\theta_2\left(\widetilde{b}_2^{CF}\right)}^{\bar{\theta}} \left[S - (1+\lambda)H(\theta)\right] dF(\theta). \end{aligned} \quad \text{(C.1)}$$

Now consider a renegotiation-proof offer $\mathbf{b} = \left(b_1^{FF}, b_2^{FF}, b_2^{CF}\right) \equiv \left(b_1^{FF}, R\right)$. By Lemma A.1, $R = \left(b_2^{FF}, b_2^{CF}\right)$ solves the following constrained maximization problem:

$$\max_{\widetilde{R}=\left(\widetilde{b}_2^{FF}, \widetilde{b}_2^{CF}\right)} SW_2\left(\mathbf{b}, \widetilde{R}\right) \quad \text{subject to (A.1).} \quad \text{(C.2)}$$

It follows from the condition $\alpha < 1+\lambda$ that (A.1) is binding for any solution to (C.2). Consequently, the f.o.c. of the optimization problem (C.2) with respect to $\widetilde{b}_2^{CF}$ at $\widetilde{b}_2^{CF} = b_2^{CF}$ is

$$\begin{aligned}
0 &= \partial SW_2\left(\mathbf{b}, \widetilde{R}\right)/\partial \widetilde{b}_2^{CF}|_{\widetilde{b}_2^{CF}=b_2^{CF}} + \eta \\
&= \theta_2'\left(b_2^{CF}\right)\left\{S-(1+\lambda)b_2^{CF} + \alpha\left[b_2^{CF} - H(\theta_2\left(b_2^{CF}\right) - e(\theta_2\left(b_2^{CF}\right))) - \psi(e(\theta_2\left(b_2^{CF}\right)))\right]\right\} f(\theta_2\left(b_2^{CF}\right)) \\
&\quad + \int_{\theta_1(\mathbf{b})}^{\theta_2\left(b_2^{CF}\right)} (\alpha-1-\lambda) f(\theta)d\theta - \theta_2'\left(b_2^{CF}\right)[S-(1+\lambda)H(\theta_2\left(b_2^{CF}\right))]f(\theta_2\left(b_2^{CF}\right)) + \eta \\
&= \theta_2'\left(b_2^{CF}\right)(1+\lambda)[H(\theta_2\left(b_2^{CF}\right)) - b_2^{CF}]f(\theta_2\left(b_2^{CF}\right)) + \int_{\theta_1(\mathbf{b})}^{\theta_2\left(b_2^{CF}\right)} (\alpha-1-\lambda)f(\theta)d\theta + \eta, \qquad \text{(C.3)}
\end{aligned}$$

where $\eta > 0$ is the Lagrange multiplier corresponding to the binding constraint (A.1).

It also holds that $b_2^{CF} = H\left(\theta_2\left(b_2^{CF}\right) - e\left(\theta_2\left(b_2^{CF}\right)\right)\right) + \psi\left(e\left(\theta_2\left(b_2^{CF}\right)\right)\right)$, of which taking derivative w.r.t. b_2^{CF} on both sides and rearranging terms yield

$$\begin{aligned}
\theta_2'\left(b_2^{CF}\right) &= \Big\{H'\left(\theta_2\left(b_2^{CF}\right) - e\left(\theta_2\left(b_2^{CF}\right)\right)\right) \\
&\quad - \left[H'\left(\theta_2\left(b_2^{CF}\right) - e\left(\theta_2\left(b_2^{CF}\right)\right)\right) - \psi'\left(e\left(\theta_2\left(b_2^{CF}\right)\right)\right)\right] e'\left(\theta_2\left(b_2^{CF}\right)\right)\Big\}^{-1} \\
&= \frac{1}{H'\left(\theta_2\left(b_2^{CF}\right) - e\left(\theta_2\left(b_2^{CF}\right)\right)\right)}, \qquad \text{(C.4)}
\end{aligned}$$

where the second equality follows from the f.o.c. regarding the optimal effort, specifically, $H'\left(\theta_2\left(b_2^{CF}\right) - e\left(\theta_2\left(b_2^{CF}\right)\right)\right) - \psi'\left(e\left(\theta_2\left(b_2^{CF}\right)\right)\right) = 0$.

Taking the fact that $\eta > 0$, substituting (C.4) into (C.3) and rearranging terms yield

$$\left(1-\frac{\alpha}{1+\lambda}\right)\left[F\left(\theta_2\left(b_2^{CF}\right)\right) - F\left(\theta_1\left(\mathbf{b}\right)\right)\right] \geq \frac{\left[H\left(\theta_2\left(b_2^{CF}\right)\right) - b_2^{CF}\right] f\left(\theta_2\left(b_2^{CF}\right)\right)}{H'\left(\theta_2\left(b_2^{CF}\right) - e\left(\theta_2\left(b_2^{CF}\right)\right)\right)},$$

which completes the proof of Lemma A.2. ■

Proof of Lemma A.3. Without loss of generality, set $\varpi_1 = 0$ and $\varpi_2 = 1$. So the binary variable W itself becomes a 0 and 1 dummy variable. Consequently, $w_{f_i} = \mathbf{1}\left(w_{f_i} = \varpi_2\right)$.

Define

$$\begin{aligned}
Q_{n_f}\left(q_1, q_2, \tau\right) &\equiv \frac{1}{n_f}\sum_{i=1}^{n_f} \rho_\tau\left(c_{f_i} - q_1 - q_2 \cdot w_{f_i}\right), \\
Q\left(q_1, q_2, \tau\right) &\equiv \mathbb{E}\left[\rho_\tau\left(c_{f_i} - q_1 - q_2 \cdot w_{f_i}\right)\right].
\end{aligned}$$

In addition, define $q(\tau) = \left(q_1(\tau), q_2(\tau)\right)' \equiv \underset{q=(q_1,q_2)'}{\operatorname{argmin}}\ Q\left(q_1, q_2, \tau\right)$. It holds that

$$\widehat{q}(\tau) = \left(\widehat{q}_1(\tau), \widehat{q}_2(\tau)\right)' = \underset{q=(q_1,q_2)'}{\operatorname{argmin}}\ Q_{n_f}\left(q_1, q_2, \tau\right). \qquad \text{(C.5)}$$

Note that all conditions required by Theorem 3 in Angrist et al. (2006) hold in our setting, with c_i in place of their Y_i, and w_i in place of their X_i. Specifically, Conditions (i), (ii) of Theorem 3 in Angrist et al. (2006) are guaranteed by Assumption 3.1 (i), (iii). And Conditions (iii), (iv) of Theorem 3 in Angrist et al. (2006) hold in our setting due to the boundedness of the support for $C|W$ and the fact that w_i is a binary variable. Moreover, since the dependent variable w_i is binary, the linear quantile regressions we consider is correctly specified. Therefore, the second result of Theorem 3 in Angrist et al. (2006), specified by their Equation (15), is applicable to our setting. Accordingly, it holds that

$$J(\tau)\sqrt{n_f}\left(\widehat{q}(\tau)-q(\tau)\right) \quad \xrightarrow{\mathcal{L}} \quad z(\tau) \tag{C.6}$$

with $J(\tau)=\mathbb{E}\left[f_{C|W}\left(q_1(\tau)+q_2(\tau)W|W\right)W\right]$, where $z(\cdot)$ is a zero mean tight Gaussian process on $L^{\infty}[0,1]$, and the weak convergence "$\xrightarrow{\mathcal{L}}$" is uniform in the sense of Chapter 1.3 in van der Vaart and Wellner (1996).

Consequently, it follows from Theorem 1.11.1 (Extended continuous mapping) in van der Vaart and Wellner (1996) that

$$\begin{aligned}
\sqrt{n_f}\left[\widehat{C}(\tau;\varpi_1)-C(\tau;\varpi_1)\right] &= \sqrt{n_f}\left(\widehat{q}_1(\tau)-q_1(\tau)\right)\\
&= (1,0)\sqrt{n_f}\left(\widehat{q}(\tau)-q(\tau)\right)\\
&\xrightarrow{\mathcal{L}} (1,0)\,z(\tau)/J(\tau)\,, \qquad\qquad\qquad\qquad\qquad\qquad\qquad (\text{C.7})\\
\sqrt{n_f}\left[\widehat{C}(\tau p;\varpi_2)-C(\tau p;\varpi_2)\right] &= \sqrt{n_f}\left[\left(\widehat{q}_1(\tau p)+\widehat{q}_2(\tau p)\right)-\left(q_1(\tau p)+q_2(\tau p)\right)\right]\\
&= (1,1)\sqrt{n_f}\left(\widehat{q}(\tau p)-q(\tau p)\right)\\
&\xrightarrow{\mathcal{L}} (1,1)\,z(\tau p)/J(\tau p)\,, \qquad\qquad\qquad\qquad\qquad\qquad (\text{C.8})
\end{aligned}$$

for τ on $[0,1]$. Again, according to Theorem 1.11.1 in van der Vaart and Wellner (1996), it follows from (C.7) and (C.8) that

$$\begin{aligned}
&\sqrt{n_f}\left\{\left[\widehat{C}(\tau;\varpi_1)-\widehat{C}(\tau p;\varpi_2)\right]-\left[\widehat{C}(\nu;\varpi_1)-\widehat{C}(\nu p;\varpi_2)\right]\right\}\\
&-\sqrt{n_f}\left\{\left[C(\tau;\varpi_1)-C(\tau p;\varpi_2)\right]-\left[C(\nu;\varpi_1)-C(\nu p;\varpi_2)\right]\right\}\\
=\;&\sqrt{n_f}\left[\widehat{C}(\tau;\varpi_1)-C(\tau;\varpi_1)\right]-\sqrt{n_f}\left[\widehat{C}(\tau p;\varpi_2)-C(\tau p;\varpi_2)\right]\\
&-\sqrt{n_f}\left[\widehat{C}(\nu;\varpi_1)-C(\nu;\varpi_1)\right]+\sqrt{n_f}\left[\widehat{C}(\nu p;\varpi_2)-C(\nu p;\varpi_2)\right]\\
\xrightarrow{\mathcal{L}}\;&(1,0)\,z(\tau)/J(\tau)-(1,1)\,z(\tau p)/J(\tau p)-(1,0)\,z(\nu)/J(\nu)+(1,1)\,z(\nu p)/J(\nu p)
\end{aligned}$$

for τ on $[0,1]$ and a given $\nu\in[0,1]$, which completes the proof of Lemma A.3. ■

An alternative identification result for $H(\cdot)$

Assumption C.1 *For some known $\theta_c \in [\underline{\theta}, \overline{\theta}]$, the following holds: $\psi_1'(e_c) = \psi_2'(e_c)$, $\psi_1'(e) > \psi_2'(e)$ for all $e > e_c$, and $\psi_1'(e) < \psi_2'(e)$ for all $e < e_c$, with $e_c \equiv \max\{e(\theta_c; \varpi_1), e(\theta_c; \varpi_2)\}$.*

Corollary C.1 *Consider the single-period setting. Under Assumptions 2.1-2, 4.2(i) and C.1, $H^{-1}(\cdot)$ is identified on $\left[\max_{j=1,2} c(\underline{\theta}; \varpi_j), \min_{j=1,2} c_u^j\right]$.*

Proof of Corollary C.1. The proof of Corollary C.1 is similar to that of Proposition 4.2. The main difference of identification in Corollary C.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 C.1, it can be shown that there exists a $\theta_c \in [\underline{\theta}, \theta^{1*}]$ such that $E(\tau_c/p_1; \varpi_1) = E(\tau_c/p_2; \varpi_2) = e_c$, 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 $C(\tau_c/p_1; \varpi_1) = C(\tau_c/p_2; \varpi_2)$ since the two distribution functions of $C(\cdot; \varpi_1)$ and $C(\cdot; \varpi_2)$ intersect only once at the cost quantile corresponding to $\theta(\tau_c)$. Consequently, e_c is identified. As a result, $H^{-1}(t)$ is identified up to $\theta(\tau_c)$ as

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

To identify $\theta(\tau_c)$, note that

$$H^{-1}(C_1(0)) \quad = \quad \underline{\theta} - e_1(0). \tag{C.9}$$

In addition, by following similar steps as in the proof of Proposition 4.2, it can be shown that

$$H^{-1}(C_1(0)) = \theta(\tau_c) - e_c + \sum_{k=0}^{\infty} \Delta\tilde{E}(\tau_k(C_1(0)).p_1). \tag{C.10}$$

Combining (C.9) and (C.10) yields

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

■

Identification under the LCC

Here we discuss how the identification strategies can be simplified under the LCC, for the most part. Though as we discuss below, under the LCC, more specific assumptions on the disutility functions are needed, due to a lack of variation in the optimal effort. Like in the main context of the paper, our discussion here is focused on the single-period setting.

As discussed in Section 3.2, a linear cost function $H(\theta - e) = \beta(\theta - e)$ induces a constant optimal effort under FP, regardless of θ. Thus, the two effort proxies X and Y (for identifying/estimating the distribution of effort) are no longer needed, as the distribution now degenerates to a single point (with probability mass one) e^{*j} conditioned on $W = \varpi_j$, for $j = 1, 2$. Note that e^{j*} is characterized by the f.o.c. $\psi_j'\left(e^{j*}\right) = \beta$. And the cutoff type θ^{j*} now satisfies

$$b^{j*} = \beta\left(\theta^{j*} - e^{j*}\right) + \psi_j\left(e^{j*}\right) \tag{C.11}$$

Still conditioned on $W = \varpi_j$, denote by $\bar{c}^j$ and $\underline{c}_{CR}^j$ the cost upper bound under FP and the cost lower bound under CR, respectively. We have

$$\bar{c}^j = \beta\left(\theta^{j*} - e^{j*}\right) \quad \text{and} \quad \underline{c}_{CR}^j = \beta\theta^{j*}. \tag{C.12}$$

Note that the lack of variation in optimal effort makes it impossible to identify $\left\{e^{j*}, \psi_1(\cdot), \psi_2(\cdot)\right\}$ without additional assumptions. For this reason, we normalize $\psi_1(e) = e^2$, and adopt a simple parameterization $\psi_2(e) = \kappa e^2$ for ψ_2, which satisfies Assumption 3.2, and is widely used in the literature. See, for example, Laffont and Tirole (1988), Rogerson (2003), Chu and Sappington (2007), and Battaglini (2007). These specifications on $\{\psi_1(\cdot), \psi_2(\cdot)\}$, together with (C.11) and (C.12), identify e^{1*} as

$$e^{1*} = \left(b^{1*} - \bar{c}^1\right)^{1/2}. \tag{C.13}$$

Consequently, accoding to (C.12) and (C.13), β is identified as

$$\beta = \frac{\underline{c}_{CR}^1 - \bar{c}^1}{e^{1*}} = \frac{\underline{c}_{CR}^1 - \bar{c}^1}{\left(b^{1*} - \bar{c}^1\right)^{1/2}}.$$

With β being identified, θ^{j*}, for $j = 1, 2$, are immediately identified from the second equation in (C.12) as $\theta^{j*} = \underline{c}_{CR}^1/\beta$. Consequently, $\left\{e^{2*}, \kappa\right\}$ are identified as the solution to

$$\begin{cases} \psi_2'\left(e^{2*}\right) = 2\kappa e^{2*} = \beta; \\ b^{2*} = \bar{c}^2 + \kappa\left(e^{2*}\right)^2. \end{cases}$$

In close forms, we have

$$\begin{aligned} e^{2*} &= \frac{2\left(b^{2*} - \bar{c}^2\right)}{\beta} = \frac{2\left(b^{2*} - \bar{c}^2\right)\left(b^{1*} - \bar{c}^1\right)^{1/2}}{\underline{c}_{CR}^1 - \bar{c}^1}, \\ \kappa &= \frac{\beta}{2e^{2*}} = \frac{\left(\underline{c}_{CR}^1 - \bar{c}^1\right)^2}{4\left(b^{1*} - \bar{c}^1\right)\left(b^{2*} - \bar{c}^2\right)}. \end{aligned}$$

With β and e^{j*} for $j = 1, 2$ being identified, we can recover the underlying θ associated with any observed cost c conditioned on $W = \varpi_j$ as

$$\theta = \begin{cases} c/\beta + e^{j*}, \text{ under FP;} \\ c/\beta, \text{ under CR,} \end{cases}$$

based on which the distribution of θ is identified. The identifications of α remains unchanged.

We note the followings at the end of this discussion: (i) The identification steps above are easily implementable for conducting estimation under the LCC; (ii) Following the existing literature, we specify $\psi_2(\cdot)$ as $\psi_2(e) = \kappa e^2$ under the LCC. As explained, this is due to the lack of variation in the optimal effort under the LCC. The necessity of adopting such a simplifying specification is further illustrated by the fact that identification would fail under a slightly more general specification $\psi_2(e) = \kappa_1 e + \kappa_2 e^2$. Specifically, under this alternative specification, $\left(e^{2*}, \kappa_1, \kappa_2\right)$ are not jointly identifiable. This is because these three parameters are characterized by merely two equations, which are $\beta = \kappa_1 + 2\kappa_2 e^{2*}$ (implied by the f.o.c. regarding the e^{2*}) and $b^{2*} = \bar{c}^2 + \kappa_1 e^{2*} + \kappa_2 \left(e^{2*}\right)^2$. $\blacksquare$