DISCUSSION PAPER SERIES

chaire eppp

IAE - PANTHEON-SORBONNE

EPPP DP No. 2014-04

Quality of Outsourced Services, Rent-Seeking and Contract Design. Evidence from Cleaning Contracts

Jean Beuve
Lisa Chever

Mars 2014

Chaire Economie des Partenariats Public-Privé
Institut d'Administration des Entreprises

Quality of Outsourced Services, Rent-Seeking and Contract Design. Evidence from Cleaning Contracts.

Jean Beuve [*]
Lisa Chever [†]

March 2014

Abstract

This paper estimates the impact of contract design on the enforceability of quality in outsourced public contracts. Using a four-year panel data of 102 cleaning contracts, our findings suggest that quality enforcement highly depends on the *ex ante* contracting process. Moreover, since those quality improvements are not accompanied by prices increases, our results suggest that public contract managers have significant leeways to reduce rent-seeking behaviors for standard transactions. Such findings are in line with propositions from the theoretical literature on the endogenous verifiability of quality. They also depart from previous studies on this issue by focusing on determinants of enforcement that are left to the managers' discretion.

Keywords: Outsourcing services, quality, enforcement, moral hazard, contract design, monitoring, incentives.
JEL Codes: D82, L15, L24

[*]CES, University of Paris I Pantheon Sorbonne
[†]Sorbonne Business School / Chaire EPPP. 21, rue Broca, 75005 Paris, France. E-Mail : lisa.chever@gmail.com

1 Introduction

In 2009, a public buyer operating in the field of social housing was sanctioned by the administrative court of Paris for disqualifying a candidate during an open call for tenders : arguing that this candidate was guilty for providing low level of service quality in past cleaning contracts, the public buyer decided to reject its candidacy at the pre-qualification stage of the procurement procedure. The court, seized by the dismissed private operator, has considered that the argument used to disqualify the claimant was unlawful. As a consequence, the public buyer was condemned to re-organize the call for tenders and to evaluate the candidacy of each operator, including the claimant. This judicial decision triggered an important organizational change on the public buyer's side. Confronted with a statutory ban to eliminate firms' candidacy on the basis of bad past performances, [1] the public authority decided to deal with firms' opportunism in another way, *i.e.* by improving the completeness of its contracts. However, since the transaction costs incurred to reduce contractual incompleteness can be substantial [Bajari and Tadelis, 2001], the extent to which this strategy is relevant is not immediately obvious. As it will be analysed, it highly relies on the ability of better *ex ante* contracting to improve quality enforcement and on the magnitude of such relationship.

Up to now, the literature has mainly focused on the disciplinary power of competition for the field to deter firms' opportunism in public procurement. In Europe, for instance, the public sector buyer is usually forced by regulation to use an open call for tender [European Commission, 2004]. This mechanism is most often considered as efficient, notably because the overwhelming majority of empirical works on this issue indicates that this procedure allows to reduce costs. [2] Nevertheless, its impact on quality is much more puzzling, suggesting that those costs savings might be achieved at the expense of quality.

As a first explanation, the complexity of the transaction can give rise to incomplete contracts that do not allow to perfectly align parties' interests and thus to deter firms' opportunism (see, e.g., Grossman and Hart [1986]). However, cleaning contracts refer to a rather simple activity in which performance measure is supposed to be easy and thus, allow the design of incentive schemes [Brown

1. See Spagnolo [2012] for an extensive discussion on this issue.
2. For instance, a meta-analysis conducted by the Australian Industry Commission [1996] on 203 different international studies on the effect of contracting on cost in public procurement reveals that the most frequently reported magnitudes of cost savings lies in the ranges of 10-30 percent.

and Potosky, 2005]. A second explanation lies on poor contract enforcement. Evidence indicates that public buyers do not apply penalty clauses, even when important damages occur [Spagnolo, 2012]. Previous empirical literature on this subject focuses on the enforcement costs to explain their limited use (see, *e.g.*, Girth [2012]; Coviello et al. [2013]). However, in our case, the public buyer tends to apply penalty clauses: low level of service still persists, suggesting that those clauses have a low incentive effect.

We complete previous empirical analyses on quality enforcement issues by investigating the role played by the *ex ante* contracting process in the deterrence of *ex post* opportunism. This approach is line with a theoretical literature that examines situations in which *'the verifiability of the agent's actions is endogenously determined by the principal's investments in drafting an explicit contract pertaining to the quality of the agent's output'* [Kvaloy and Olsen, 2009]. In this literature, investing in the *ex ante* contracting process can reduce contractual incompleteness and thus increase the court's ability to verify that the agent's actions are line with the intend of the contract. This may be especially efficient in moral hazard settings [Kvaloy and Olsen, 2010], where the supplier has some private information and needs some incentives to provide efforts.

To investigate the role of contractual specification on the deterrence of moral hazard, we use an original database coming from Paris Habitat-OPH, the local public buyer sanctioned by the administrative court in 2009. We have access to monthly information related to 102 allotted contracts signed with a set of private operators in the sector of house-cleaning over a four year period. Two significant dates need to be pointed out. First, in order to facilitate performances comparisons and to implement a penalty system, the public authority decided in 2004 to build a tool so as to systematically assess the level of service quality.[3] Second, in April 2010 (and in reaction to the court decision previously mentioned), the public buyer decided to implement two major changes : all the contracts launched after this date include (1) more detailed specifications (in particular, about the way to reach a satisfying quality of service) and (2) a reinforced enforcement regime both for controls and penalties. We interpret these changes as an increase of contractual completeness. As a result, we have panel data which allows us to observe the evolution of quality before and after an exogenous shock on the level of contractual completeness, leaving us with a quasi-natural experiment. Moreover, our information about the value of the penalties paid and the frequency of controls

3. This tool take the form of an evaluation grid, based on multiple and objective criteria, which should be monthly completed for each contract.

enables to disentangle the direct impact of the formal contract from its indirect impact. We indeed capture this indirect impact by investigating whether controls and penalties provide stronger incentives after the change in contract design.

Our findings suggest that bearing the costs of reducing contractual incompleteness significantly improves the delivered quality: more complete contracts not only reduce the average level of moral hazard (direct impact) but also optimize the monitoring of the contract (indirect impact). Regarding the latter aspect, we indeed find that controls and penalties have the ability to deter firms' opportunism only when contracts are well-specified. In addition, we give evidence that quality improvements are made possible without prices raising. As a consequence, our analysis illustrates how an exogenous 'bad' event (the court decision) finally leads to an unexpected and profitable organizational change. On the one side, for standard goods or services like cleaning activities, the transaction costs that enable to reduce contractual incompleteness are incurred only once. On the other side, the resulting long term reduction of firms' opportunism is not associated with a price increase while allowing to spare on the costs that have to be engaged to monitor the suppliers' performances.

We believe our results contribute to the emerging empirical literature on the issue of public contract enforcement [Girth, 2012; Coviello et al., 2013]. While previous studies investigate the impact 'exogenous' factors on public contract enforceability, we shed light on a strategy which is not only at the discretion of public managers but also in line with the theoretical literature on endogenous verifiability (see,of *e.g.*, Kvaloy and Olsen [2009, 2010]).

The rest of the paper stands as follow. Section 2 presents the conceptual framework and puts forward the hypothesis we aim to test. Section 3 then gives some details about the institutional context, the data and our empirical strategy. The results from our estimations are provided in Section 4. Finally, we discuss those results and potential recommendations for outsourcing of public services in Section 5.

2 Conceptual framework

Numerous quantitative studies have compared costs of in-house public management and costs of private management when contracts are awarded through a competitive tendering. In their overwhelming majority, those studies conclude

that outsourcing achieves reduction in government expenditures in a great variety of sectors such as refuse collection industry [Domberger et al., 1986; Dijkgraaf and Gradus, 2004], road maintenance [Blom-Hansen, 2003], vehicle and warehousing maintenance [Domberger et al., 2002], cleaning and housekeeping [Domberger et al., 2004] and [Milne and Wright, 2004] and even prisons [Cabral and Saussier, 2013]. Competition for the field thus effectively prevents rent extraction by encouraging operators to bid competitively.

Nevertheless, abandoned rent on costs can be recovered on the quality of the service (*quality-shading hypothesis*). Nowadays, although the effect of outsourcing on quality is of fundamental importance for the efficient organization of public services, empirical studies which examine quality issues exist in far fewer numbers than those on costs savings. Moreover, the few number of existing studies provides mixed evidences. While some of them suggest that service quality had been stagnant or risen when contracting out [Domberger et al., 1995, 2004; Savas, 1977], others reach the opposite conclusion of reduced quality following contracting out [Evatt Research Centre, 1990; Hartley and Huby, 1986; Ascher, 1987]. An interpretation of those mixed results relies on contract specification and monitoring: as put forward by the Australian Industry Commission [1996], quality issues are primarily a result of poor application of the process of outsourcing rather than outsourcing *per se*.

From a theoretical perspective, we can distinguish two main explanations about why the outsourcing process can entail quality issues. On the one side, a first strand of the literature focuses on the role of contractual incompleteness (see, *e.g*, Grossman and Hart [1986]): when dealing with some transactions, writing complete agreements might be prohibitively costly. The resulting contractual blanks prevent parties from safeguarding their respective interests. On the other side, even in the case of complete contracts (*i.e* contracts for which performances are measurable), *ex post* opportunistic behaviours can be due to asymmetries of information between parties [Jensen and Meckling, 1976]: because operators may not share the buyer's goals and be more familiar with the details of the task, they may have both motive and opportunity to behave in ways that maximize their own interest at the expense of the public authority. In other words, they are able to 'shirk' by delivering a lower service quality. Solutions lie on the provision of outcome- or behaviour-based contracts.

In the specific case of cleaning activity, contracting on quality is supposed

to be rather simple.[4] Consequently, contractual incompleteness should not be an issue. Moreover, those contracts include incentive schemes that should align parties' interests and limit informational rents. Despite of that, quality issues often raised: in the data we explore about cleaning activities, the persistence of quality issues are illustrated by frequently imposed penalties, users complaints, breach and early termination of contracts.

Explanations of such paradox can be found in the literature on endogenous verifiability. Indeed, the seminal papers of Townsend [1979] and Dye [1985] highlight that costly contracting and imperfect enforcement are important vehicles to understand the nature of transactional relationships. While the classic moral hazard approach assumes perfect enforcement (see, *e.g.*, Holmström [1979]) and models of incomplete contracting consider that contracting is prohibitively costly so that legal enforcement is impossible (see, *e.g.*, Grossman and Hart [1986]), more recent contributions overstep those assumptions by dealing with costly contracting and endogenous verifiability [Kvaloy and Olsen, 2009, 2010]: the time and efforts spent on the contracting process determine the level of verifiability.

Contracting on quality is indeed challenging. Even if service quality may be identified in terms of performance characteristics, their assessment may require subjective judgement rather than mere accumulation of data [Jensen and Stonecash, 2005]. For instance, in the specific case of cleaning services, the only way quality can be measured is through personal observation and what constitutes a high standard of cleanliness may vary from one observer to another [Domberger and Jensen, 1997]. In such a situation, improving contract details can help in reducing ambiguities between parties about the intends of the contract and thus, increase its enforceability. This threat might then help to overcome the service providers' temptation to reduce efforts.

4. As an illustration, Brown and Potosky [2005] sent a survey to public managers about the transaction cost dimension of a variety of basic local government services in order to build a typology of "ease of measurement" for service performances. More precisely, they ask manager to determine this measurement easiness in a five point scale, giving that "*a service is easy to measure if it is relatively straightforward to monitor the activities required to deliver the service and to identify performance measures that accurately represent the quantity and the quality of the service*". They also precise that for easy-to-measure services, "*government officials can easily write a contract and clearly specify the activities and outcomes for the vendor to perform and achieve*". This way, they identify very easy-to-measure outsourced activities, such as payroll, commercial solid waste collection and street and house cleaning (score < 2) and very difficult-to-measure ones, such as child welfare programs, drug and alcohol treatment and operation of mental health program (score > 4). According to such a typology, contracts established to outsource cleaning services might be rather complete and quality-shading might be rather scarce on this sector.

The data currently available to us in this study are closed to the framework of Kvaloy and Olsen [2010] where contracting is costly and enforcement is probabilistic (the probability that the incentive contract will be enforced by a court of law is determined by the costs spent on contracting). In their paper, better contract specification leads to higher-powered incentives. Nevertheless, as far as we know, there is no hard empirical proofs of such statements. This is precisely how this paper contributes to the literature by filling this gap.

3 Data and empirical strategy

3.1 Institutional framework

Over the last decades, outsourcing activities to external providers has become a fairly common practice for governments. As a consequence, the way it influences the costs of public services is the focus of academics' and practitioners' interests. The outsourcing of public activities is regulated by European directives, which are then transposed into national law through the 'French Public Procurement Code'. As soon as service contracts reach the EU-thresholds (around 200 k€), both regulations constrain the public buyers to use the traditional open call for tenders. [5] This procurement process is made of different steps. First, the buyer defines its needs and it launches a publicity. Second, firms send documents related to their candidacy (their references, their number of employees, their competences, etc.) and their offer. Third, the buyer analyses the different bids. If they are satisfying, the associated offers are also analysed. Finally, the winner is the "most economically advantageous offer". This selection process is deliberately rigid so as to respect the principles of the directive: equal treatment, non-discrimination, mutual recognition, proportionality and transparency (European Commission [2004]). In theory, this rigidity is supposed to ensure the efficiency of competitive mechanisms because any firm can submit a bid which is evaluated according objective and transparent criteria; moreover, ignoring the number and the offers of rivals incite bidders to reveal their private information. Nevertheless, this statement is true only if, in particular, it is possible to contract on and monitor the quality of the service. Otherwise, the *ex ante* competition does not prevent opportunistic

5. While a derogation is possible in specific cases, it is still the most used mechanism: according to a EU-report, 68% of service contracts awarded between 2006 and 2010 followed an open procedure (see the 'Impact of the Effectiveness of EU Public Procurement Legislation', p.12, Part 1, European Commission (2011)).

behaviors and firms can decrease the delivered quality.

3.2 Cleaning contracts

In our set of house-cleaning contracts, contract specification entails a precise description of quality standards, in terms of direct service provision (detailed description of the tasks, the resources to be used and the calendar of interventions) as well as in terms of relational aspects (compliance with deadlines when answering to the buyer's requirements and when providing contractual documents such as service manuals and periodical activities' reports). Concerning the monitoring, the contract specifies how the quality will be evaluated and by whom.

As previously mentioned, the public buyer we study implemented a detailed evaluation grid in order to minimize the issue of subjective judgement and to allow comparisons based upon an objective quality-scoring identically constructed across all cleaning contracts. Evaluation are made during 'contradictory controls', *i.e.* controls in the presence of the representative of the public buyer and the responsible for technical and administrative matters of the cleaning company, organized once per month. The contract holder is notified 48 hours in advance. [6]

The completion of the evaluation grid (called *quality control sheet*) leads to a final mark out of 100. This scoring then allows to contractually defined incentives mechanisms. Indeed, if the obtained mark is less than 80, a second contradictory control is planned 48 hours later and penalties are imposed to the contract holder. Those penalties increase if the mark obtained during the second control is still less than 80. The goal of such penalties is to provide coercive means at the disposal of the public buyer to enforce contractual specifications and, in case, to punish any breach of contractual obligations. Moreover, the public buyer is contractually authorized to impose penalties when the cleaning companies fail to fulfil their obligations in terms of mandatory documents supply. Finally, contracts also contain a cancellation clause that can be applied in the event of repeated failures and/or when accumulated penalties exceed a contractually predetermined threshold.

Regarding such contractual arrangements, one could imagine that the public

6. It is important to note that such a delay does not enable the private operator to react by 'falsifying' the quality evaluation through a short-term effort just before the control. Indeed, a lot is composed by between 544 and 3.066 accommodations and controls are ran for random buildings. Furthermore, cleanliness is mainly derived from the length of effort over time and a short and intense burst of activity might be not sufficient to obtain a good evaluation.

buyer we study is well protected against opportunism. This is not what we observe in the data. As previously said, frequently imposed penalties, users' complaints, breach and early termination of contracts persist despite the use of open auctions, the definition of quality standards and the monitoring of the contract.

To tackle those problems, the public buyer has decided to modify its formal contracts in three directions. From the specification point of view, the new versions of the '*General Conditions*' and of the '*Guide of Special Techniques Specifications*' are more precise regarding the obligations of the operator. For instance, the contract describes more technically the way cleaning must be performed and includes a glossary of cleaning operations. From the monitoring point of view, the new contractual arrangement increases the level of details provided to the cleaner about how the evaluation is made and adds the possibility of unplanned and not contradictory controls by public agents. Finally, concerning incentives, the new contract adds new categories of penalty clauses and increases their amounts. Aside this formal transformation, the public buyer also decided to be more rigorous in the application of penalties. Table 2 (in the appendix) highlights all the differences between old (launched before April 2010) and new contracts (launched after April 2010).

3.3 Data

The public buyer we study organizes its cleaning activity by establishing a geographic allotment. Indeed, the buildings located in a same area correspond to a given lot j whose characteristics remain rather constant over time. It allows us to follow a lot over time and to assess whether different types of contract are associated with different quality levels. Over the period of 48 months we study (between July 2008 and June 2012), our unbalanced panel database [7] of 49 lots experimented a total of 102 contracts. More precisely, 45 lots have been renewed one time, [8] whereas 4 have been renewed two times. We know that these contracts are shared out among 13 firms and are managed by 6 different departments on the buyer's side (each department is in charge of a geographical area). We also have information regarding the monthly level of implemented quality, the controls frequency, the applied penalties, the tendering phase and the price of the contract.

7. The panel database is unbalanced because some quality indicators are missing. Additional details are given when presenting the data. Additional econometric specifications are also run to deal with the biases this issue might generate.

8. It means that we observe a lot through two different contracts and sometimes through two different suppliers

Summary statistics of our variables are presented in Table 1, whereas Table 4 (in the appendix) provides descriptive statistics when distinguishing variables before and after the change. We call panel A the 50 contracts launched before April 2010 and Panel B the 52 launched after this date.

3.3.1 Dependent variable

We are mainly interested in explaining the variable $Quality_{ijt}$. It measures the level of quality delivered by firm i on lot j at time t. As previously said, quality indicators are based on a scale of 0 to 100. [9] The average level of quality is 88.91, which is around 9 points beyond the threshold implying penalties. Quality scores of panel B are significantly higher of two points than panel A ($p < 0.01$).

Figure 1 reports the average value of quality at each period and shows that $Quality_{ijt}$ significantly increased after April 2010. Since most of the on-going contracts belong to panel B after April 2010, this observation strongly corroborates our intuition that the level of service quality increased after the modifications of the contract design.

3.3.2 Main independent variable

We want to assess the impact of contractual completeness on moral hazard. The precise information we have enables to disentangle a direct impact due to the new contract design from an indirect impact due to more efficient controls and penalties.

Contract design

We first constructed the variable $NewDesign_{jt}$, which is a binary variable taking the value 1 if the contract running on lot j at time t belongs to Panel B, 0 otherwise. This variable captures the exogenous shock affecting all the contracts launched from April 2010. As previously said, it resulted in specifying the contracts in more detail, in reinforcing the penalties clauses and in including a new procedure of control. We interpret those changes as an increase of contractual completeness. According to the literature on the endogenous verifiability, such change may increase contract enforceability. We may then expect a positive impact of the variable $NewDesign_{jt}$ on the level of service.

9. This mark is communicated to the supplier, but is not a public information.

Controls

The variable $ControlFreq_{ijt-1}$ corresponds to the number of times the quality delivered by firm i on lot j at time t has been controlled, divided by the maximal number of times it could have been controlled. On average, it is equals to 0.54. It is higher in Panel A than in Panel B. It reflects that our buyer decreased controls frequency while reinforcing contractual requirements and imposing higher penalties. One interpretation would be a substitution effect between the greater amounts of penalties and the necessity to carry out assiduous controls. In other words, if the expanded threat of punishment disciplines firms, then regular controls are less useful. We expect that a firm managing a frequently monitored contract (*i.e.* having a larger $ControlFreq_{ijt-1}$) feels more intensely under the scrutiny of the buyer, leading us to anticipate a positive impact of the variable $ControlFreq_{ijt-1}$ on our dependent variable.

Incentives

$Penalties_{ijt-1}$ is the second variable that accounts for contract monitoring. It corresponds to the full amount of penalties paid by firm i on lot j at time t. Around 731 additional euros per contract have been paid. When considering the contracts' size,[10] this amount is rather low and reflects that penalties are a least resort. Nevertheless, a significant difference of 150 euros of penalties distinguishes Panel A from Panel B, which sheds some light on our buyer's decision to reinforce the sanction mechanisms. This evolution appears even more clearly in Table 3 that summarizes the value of penalties paid each year by the cleaning contracts' suppliers. Since penalties account for small amounts and since reaching a fixed amount of penalties lead to a breach of the contract, we claim that it is the accumulated value of paid penalties that may have an incentive effect on the suppliers' decision to improve their current performances. We thus expect a positive impact of the variable $Penalties_{ijt-1}$ on $Quality_{ijt}$.

The variables $ControlFreq_{ijt-1}$ and $Penalties_{ijt-1}$ capture some heterogeneity regarding the contract monitoring. However, when the contract enforceability increases, we expect that monitoring firms' outcomes provides stronger incentive to deliver high quality. Therefore, $ControlFreq_{ijt-1}$ and $Penalties_{ijt-1}$ may have a larger impact after the change in contract design.

10. The final price of the contracts we study are made of two parts: a fixed part and a variable part; this last part depends on some buyer's needs which are unknown at the awarding stage. We only have precise information about the winning bid for the fixed part. It accounts for at least 70% of the final price and equals around 395 642 k€.

3.3.3 Control variables

We use additional controls to capture heterogeneity across contracts. Indeed, some dimensions like the degree of competition, the price or the scope of the contract can affect the level of delivered quality. The variables associated with these dimensions and their expected impacts on quality are described below.

Geographical allotment

Our variable $NbAccomodations_{jt}$ captures the number of accommodations which are included in lot j at time t. We indeed aim at exploiting the panel nature of our data by following a same lot over time. To correctly perform it, the lots have to remain stable. Nevertheless, we cannot rule out that our buyer will build, buy or sell buildings, therefore affecting the characteristics of the lots. The variable $NbAccomodations_{jt}$ enables to account those types of changes. Since our buyer allots its cleaning activity because he expects larger lots to be more difficult to manage, we suspect that the number of accommodations negatively impacts on the level of delivered quality.

Competition

The variable $NbOffers_{jt}$ stands for the number of offers received by the buyer for lot j at time t. An increase in competition is supposed to be beneficial: it incites firms to reveal their private information and to lower their rents. However, more competition might also encourage aggressive bids at the expense of quality. Indeed, as previously underlined, asymmetries of information might allow firms to shirk on quality during the execution of the contract. Therefore, the impact of competition on quality in not consensual and hard to make out.

An additional difficulty is that the degree of competition is often considered as an endogenous variable, *i.e.* a variable varying for unobserved reasons that also affect outcome variables (like quality).[11] So as to clarify the nature of the variable $NbOffers_{jt}$, Figure 2 describes the relationship between the number of offers and the year the contract is awarded. All the contracts starting from 2010 include the new design. We can see that the number of offers tends to increase in 2010 and 2011. Although it seems surprising that strengthening the contract design generates a boom in the number of potential suppliers, practitioners argue that this change sent to firms the signal that the buyer was unsatisfied with its current main suppliers and aimed at finding new performing firms. This signal may have renewed the set of interested firms, creating a chock on the degree of

11. See, e.g., Coviello and Mario Mariniello [2012] or Amaral et al. [2013]

competition. If this increase in competition actually enabled to renew the set of suppliers (19% of the contracts launched in 2010 are awarded to new firms) by attracting performing firms, adverse selection might have been reduced. However, if more competition ended up in encouraging aggressive bids, it may have resulted in damaging quality. Finally, the variable $NbOffers_{jt}$ is problematic for two reasons: its impact on quality is hard to anticipate and we don't know whether we should consider it as an independent variable.

Prices

We finally built the variable $Price_{ijt}$ which corresponds to the bid of the winning firm i for lot j at time t divided by the number of accommodations. This variable imperfectly captures the competitiveness of the winning offers. [12] Although bids competitiveness might signal firms' efficiency, low prices might also reflect a strategic or a naive behaviour of candidates. [13]

In addition, for the same grounds as those put forward when presenting the variable $NbOffers_{jt}$, the variable $Price_{ijt}$ is likely to be endogenous. Once again, we try to clarify it by analysing its evolution over time (see Figure 2). We can see that winning bids tend to increase from 2010. Thus, it seems that the new contract design is associated with less competitive bids: firms would compensate the costs they incur to fulfil the increasing expectations of the buyer by posting higher prices. Even though the potential trade-off between price and quality is a matter of concern in this paper, considering $Price_{ijt}$ as an independent variable that impacts on the delivered quality may be an issue.

Outcome variables

We have to decide about whether introducing $NbOffers_{jt}$ and $Price_{ijt}$ as control variables. On the subject of 'bad controls', Angrist and Pischke [2008] (p.64, chap. 3) give the following reasoning:

"*Some variables are bad controls and should not be included in a regression model even when their inclusion might be expected to change the short regression coefficients. Bad controls can be defined as variables that are themselves outcome variables in the notional experiment at hand. Consequently, bad controls might just as well be dependent variables too. On the other hand, good controls are variables that we can think of as having been fixed at the time the regressor of interest was determined*".

12. As previously mentioned, the final price of the contract are made of a fixed part and a variable part. The variable $Price_{ijt}$ is built thanks to information we have about the winning bid for the fixed part, which accounts for at least 70% of the final price.

13. 'Naive behaviour' refers to the winner's curse issue (see, *e.g.*, Hong and Shum [2002]).

On the one hand, the price and the degree of competition may have been affected by the change in contract design. They could be considered as outcome variables. On the other hand, $NbOffers_{jt}$ and $Price_{ijt}$ are fixed before the quality delivering. We could thus decide to use them as control variables. However, we can reasonably assume that prices and (at least, part of) the quality are simultaneously determined by the agent at the awarding stage. We do not have this problem with the number of offers: it is fixed before the delivering of quality and it results from rival's decisions. As a consequence, we decide to use $NbOffers_{jt}$ as a control variable and $Price_{ijt}$ as an outcome variable; we separately explore the determinants of the latter variable in a second step of our analysis.

3.4 Models specifications

We are interested in assessing whether better specified contracts enable to improve the level od quality. We can investigate this question because the buyer we study built the quality indicator $Quality_{ijt}$. Therefore, we first estimate the following model (1):

$$Quality_{ijt}=\beta_1+\beta_2.NewDesign_{jt}+\beta_3.Y_{jt}+W_j+\epsilon_{ijt} \quad (1)$$

where $NewDesign_{jt}$ is our first main variable of interest capturing the change in the formal contract, Y_{jt} is a vector of variables capturing the characteristics of the lot j at time t. We abstract unobservable biases due to the nature of the lots by adding lot fixed effects (W_j). This first model is a simple "before-after" estimation: it assesses the impact of the switching from panel A to panel B. Since our main goal is to investigate the impact of new contracts on moral hazard issues, we second run a model with firm fixed effects, Z_i:

$$Quality_{ijt}=\beta_1+\beta_2.NewDesign_{jt}+\beta_3.Y_{jt}+W_j+Z_i+\epsilon_{ijt} \quad (2)$$

Equation (2) enables to have a more conservative estimation regarding the impact of the new contract design on moral hazard issues. Indeed, if the coefficient associated to the variable $NewDesign_{jt}$ decreases when switching from equation (1) to equation (2), it means that part of the increase in quality comes from the selection of more efficient firms. In equation (2), the variable $NewDesign_{jt}$ only captures some changes in the level of quality that are related to moral hazard

issues. Then, we can disentangle the impact of the formal contract itself from the frequency of the controls and the penalties by running a third model with the vector X_{ijt}=($ControlFreq_{ijt-1}$, $Penalties_{ijt-1}$):

$$Quality_{ijt}=\beta_1+\beta_2.NewDesign_{jt}+\beta_3.Y_{jt}+\beta_4.X_{ijt-1}+W_j+Z_i+\epsilon_{ijt} \quad (3)$$

Finally, we also run a fourth model to test whether the formal contract determines the efficiency of the controls and the applied penalties. To do so, we add an interaction term between the variable $NewDesign_{jt}$ and the demeaned value of the variables related to contract enforcement ($\widehat{X}_{ijt} = X_{ijt}$-$\bar{X}_{ijt}$).[14] We obtain the last following equation:

$$\begin{aligned} Quality_{ijt}=&\beta_1+\beta_2.NewDesign_{jt}+\beta_3.Y_{jt}+\beta_4.X_{ijt-1} \\ &+\beta_5.(NewDesign_{jt}*\widehat{X}_{ijt-1})+W_j+Z_i+\epsilon_{ijt} \quad (4) \end{aligned}$$

In this model, while the vector of coefficients β_4 captures the influence of controls and penalties before the changes in contract design, β_5 captures their influence after the change. Moreover, demeaning the variables X_{ijt} in the interaction term enables to assess the marginal impact of penalties and controls after the change in contract design.

4 Results

4.1 The determinants of quality

Table 6 exhibits our baseline results for the effect of the contractual completeness on quality enforcement. In Model 1, which is a simple before-after, we find that the change in contract design has a significant and positive impact on quality. When switching from Model 1 to Model 2, we add firms fixed effects. The coefficient associated to the variable $NewDesign_{jt}$ slightly decreases. We interpret it to be a sign that the reduction of moral hazard issue is the main driver of the quality improvement.

14. See pages 68-69 of chapter 4 in Wooldridge [2001] for the explanation regarding the construction of the interaction term.

We then add the variables $ControlFreq_{ijt-1}$ and $Penalties_{ijt-1}$ in Model 3. The coefficient associated with the variable $NewDesign_{jt}$ remains positive and significant. It corroborates the central idea of the paper: improving the contract details reduces the suppliers' temptation to shirk on quality. In other words, the formal contract itself acts as a discriminating devise.

However, in Model 3, we do not find that the variables related to the *ex post* monitoring influence the level of quality. By distinguishing their respective impact before and after the change in contract design, model 4 allows to further investigate this result. In model 4, we indeed see that the formal contract itself has a sizeable influence on the efficiency of the *ex post* monitoring : our specification exhibits a negative effect of $ControlFreq_{ijt-1}$ and $Penalties_{ijt-1}$ before the change in contract design, whereas their effect becomes positive and significant after the change in contract design.

In addition, although the degree of competition does not influence the delivered quality, we unsurprisingly find that smaller lots are associated to higher performances.

4.2 Robustness checks

4.2.1 Testing selection bias

Quality control should be done on a monthly basis but the examination of the data reveals that controls are not always performed. As a consequence, our dataset suffers from missing information corresponding to situations in which controls were not made. Since the decision to make a control is highly decentralized, left to the caretaker's discretion, whose motives are hard to make out, the reasons we may invoke to justify this sample selection are multiple. Therefore, the way this sample selection affects our estimates is difficult to anticipate.

The variable $Observe_{ijt}$ is a dummy variable: it takes the value one if the quality is controlled, 0 otherwise. It indicates that $Quality_{ijt}$ was not measured in 39% of all cases. This sample selection is important and could therefore bias our estimates. In order to tackle this issue, we can use a two-step heckman method [Heckman, 1979]. Provided that we achieve to explain why quality is observed or not, it enables to extrapolate the missing quality indicators as if they would have been observed. Thereafter, the model indicates whether the bias due to sample selection is severe and it accounts for the bias effect both on the dependent and

on the independent variables.

The first step of the procedure corresponds to the selection equation which models the probability of being observed. The second steps corresponds to the corrected outcome equation: it explains the level of quality, given it is observed. However, to be over-identified, the model requires at least one instrument to be included in the first step but not in the second step. This variable must be correlated with the variable $Observe_{ijt}$ (instrument relevance condition), but not with any unobservable that could influence the variable $Quality_{ijt}$ (exclusion restriction condition).

We suspect $ControlFreqOthers_{ijt-1}$ to be a valid instrument. It measures the rate of control at (t-1) on a sub-sample of observations. This sub-sample is made of all the observations related to the period (t-1) with the exception of both the observations related to the firm i and the observations related to the department managing lot j.[15] Regarding the relevance condition, we assume that people having both the same job and the same employer may observe and influence each others, through a spillover effect. This way, we overcome the lack of clarity regarding caretaker's motives to perform controls by considering that, on average, they may share the same motives on the short term: one caretaker's diligence to carry out a control depends on the observation of the others caretaker's diligence, which is captured through their control rate at (*t-1*). So as to respect the exclusion restriction, we build our instrument on a sub-sample of observations: we exclude those which are likely to be correlated with the unexplained performances of firm i on lot j at time t. Thus, the sub-sample excludes the observations related to firm i and those related to the department managing lot j. Indeed, a firm is likely to observe the past controls frequency of its territorial department and/or the past controls frequency of its other ongoing contracts: this information may influence its behaviour.[16] On the contrary, this firms should not be aware, at least in the short term, of the caretakers' propensity to perform controls in the other departments regarding the other firms: this is true assuming that firms do not communicate among each other on a highly regular basis. As a consequence, this private information of the caretaker should not influence a firm's incentives toward quality. Moreover, this private information should also not influence the

15. To calculate the value of the variable $ControlFreqOthers_{ijt-1}$, we look at the total number of control performed at *t-1* (in the other departments and with the other firms) and we divide this number by the total number of controls that could have been performed at *t-1* (in the other departments and with the other firms).

16. For instance, a firm may be more prone to increase the quality of service if it observes that the caretakers tend to increase their controls.

firms' rating which is based on objective criteria.

Results of the two-step Heckman estimates are presented in Table 7 in the appendix. Whatever the specification, we can see that our instrument is significant ($p < 0.01$), satisfying the relevance condition. The variable *Lambda* captures the impact of the first stage on the second stage. Given that *Lambda* is negative and sometimes significant, we conclude that the probability of being observed is higher when the quality is lower, validating the existence of a sample selection bias. However, our main findings remain extremely stables. There is one notable change in model 8: the coefficient associated with the variable $NewDesign$ slightly decreases. Since this model does not seem to be significantly affected by the sample selection (*Lambda* is no more significant), model 4 might be more relevant than model 8 to estimate equation (4).

4.2.2 Testing shock exogeneity

As previously mentioned, the modification of the contract design by the public buyer is not an *ex nihilo* decision. On the contrary, it is a reaction to a judicial decision of the administrative court of Paris in June 2009 (see Figure 3 that reports the timing of the events).

If ever this court' s decision impacted on the buyer's and/or firms' behaviors, we may fear of not having a quasi-natural experiment. To check that, we replicate our estimates by including the variable $Decision_t$ which is a dummy variable taking the value 1 after May 2009. Results are presented in Table 8 in the appendix. Whatever the specification, we can see that the decision has no impact neither on the buyer's propensity to observe the quality, nor on the level of delivered quality: parties did not adapt their behavior to this decision. Moreover our results still remain perfectly stable and confirm that the average level of quality clearly depends on the change in contract design and its enforcement, not on the decision of the court. It corroborates what we observe on Figure 1.

4.3 Prices and rent-seeking

To complement our analysis, we explore the determinants of prices. In particular, we want to assess whether the new contract design had a significant impact on prices, *i.e.* whether the public buyer has to arbitrate between price and quality. Therefore, we estimate the following equation:

$$Price_{ijt}=\beta_1+\beta_2.NewDesign_{jt}+\beta_3.PriceIndex_t+\beta_4.Y_{jt}+W_j+Z_i+\epsilon_{ijt}$$

This estimation allows to assess the impact of the change in contract design, given the nature of the lot (we add the lot fixed effects, W_j) and the degree of competition. We also aim at disentangling the impact of the identity of the supplier i from the impact of the change in contract design. However, since we run our estimations on a (rather small) sample of 102 contracts, we cannot simultaneously add lot and firm fixed effects. We address this issue by testing various specifications: each one includes a specific category of fixed effects. This strategy still allows to test whether results are stable across specifications.

Since the contracts we study have been awarded between 2004 and 2011, we are likely to observe a price increase, not because of the new contract design, but because of a general price increase in the sector (which is partially collinear to the variable $NewDesign_{jt}$). To address it, we also include a price index of cleaning services: we collected it on the website of the French National Institute of Statistics (the "INSEE") and call it $PriceIndex_t$.

Results of estimates are presented in Table 9. The number of offers has the expected impact and is line with previous results from the literature that more competition attracts lower bids. Moreover, we find that the new design has no significant impact on the received prices, whatever the specification we consider. It highlights the previous existence of rent-seeking behaviors from cleaning operators which actually have sufficient leeway to increase quality without raising bids.

5 Conclusions

The results we obtain in this study show that reducing contractual incompleteness enables to significantly reduce moral hazard issues, through direct and indirect effects. Indeed, while more complete contracts are associated with lower moral hazard, they also enable to increase the incentive power of controls and penalties processes.

In addition, we find that this improvement does not necessarily result in a significant increase in prices. It validates the idea that asymmetries of information result in rent-seeking behaviours that can be diminished by providing the good incentives.

The solution we provide is appropriated for standard transactions. However, for single-use contracts, bearing the costs of reducing contractual incompleteness might be irrelevant. Consequently, this paper also has important message concerning the way outsourcing public services is organized in the European Union. As illustrated by the previously mentioned decision of the administrative court of Paris, European rules in public procurement do not allow to take past experiences and reputation into account. While this obligation increases transparency and thus, limits abuses in discretion with public funds, it still appears as being insufficient to systematically obtain the best value for money. Drawbacks come from the fact that those rules only put the emphasis on the awarding process, which ensures, under rarely gathered conditions, an efficient contract execution. In the end, when awarding custom made contracts, public managers still have to find a way to address the issue of contractual incompleteness and contract enforcement.

6 Acknowledgments

The authors are grateful to Jennifer Arlen, Axel Gauthier, Ricard Gil and Stephan Manning. We also thank conference participants - Journée de la Microéconomie Appliquée, June 2013; International Conference 'Contracts, Procurement and Public-Private Arrangements', June 2013; International Society on New Institutional Economics, June 2013; European Group for Organization Society Conference, July 2013 - for their comments and questions on preliminary versions of this paper.

References

Amaral, M., Saussier, S., and Yvrande-Billon, A. (2013). Expected number of bidders and winning bids: Evidence from the london bus tendering model. *Journal of Transport, Economics and Policy*, 47(1):17–34.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Ascher, K. (1987). *The Politics of Privatisation: Contracting Out Public Services*. MacMillan Education, London.

Australian Industry Commission (1996). *Competitive Tendering and Contracting by Public Sector Agencies*. Australian Government Publishing Service.

Bajari, P. and Tadelis, S. (2001). Incentives versus transaction costs: A theory of procurement contracts. *RAND Journal of Economics*, 32(3):387–407.

Blom-Hansen, J. (2003). Is private delivery of public services really cheaper? evidence from public road maintenance in denmark. *Public Choice*, 115(3-4):419–38.

Brown, T. L. and Potosky, M. (2005). Transaction costs and contracting: The practitioner perspective. *Public Performance and Management Review*, 28(3):326–351.

Cabral, S. and Saussier, S. (2013). Organizing prisons through public-private partnerships: a cross-country investigation. *Brazilian Administration Review*, 10(1):100–120.

Coviello, D. and Mario Mariniello, M. (2012). Publicity requirements in public procurement: Evidence from a regression discontinuity design. *Working Paper*.

Coviello, D., Moretti, L., Spagnolo, G., and Valbonesi, P. (2013). Court efficiency and procurement performance. *Working Paper*.

Dijkgraaf, E. and Gradus, R. (2004). Cost savings of contracting out refuse collection. Technical report, EconWPA.

Domberger, S., Hall, C., and Jeffries, M. (1995). *Competitive Tendering and Contracting in Commonwealth Government Agencies: The 1995 Survey Findings*. CTC Research Team, Graduate School of Business, University of Sydney.

Domberger, S., Hall, C., and Skinner, V. (2004). *Contracting of Services in the Western Australian Public Sector: The 1995 Survey Findings*. Graduate School of Business, University of Sydney.

Domberger, S. and Jensen, P. (1997). Gaining from outsourcing. *New Economy*, 4(3):159–163.

Domberger, S., Jensen, P., and Stonecash, R. (2002). Examining the magnitude and sources of cost savings associated with outsourcing. *Public Performance and Management Review*, 26(2):148–168.

Domberger, S., Meadowcroft, S., and Thompson, D. J. (1986). Competitive tendering and efficiency: the case of refuse collection. *Fiscal Studies*, 7(4):69–87.

Dye, R. A. (1985). Costly contract contingencies. *International Economic review*, 26(1):233–250.

European Commission (2004). Directive 2004/18/ec of the european parliamant and of the council.

Evatt Research Centre (1990). *Breach of Contract: Privatisation and the Management of Australian Local Government*. Sydney, Pluto Press.

Girth, A. M. (2012). A closer look at contract accountability: Exploring the determinants of sanctions for unsatisfactory contract performance. *Journal of Public Administration Research and Theory*.

Grossman, S. J. and Hart, O. (1986). The cost and benefits of ownership: A theory of vertical and lateral intégration. *Journal of Political Economy*, 94(4):691–719.

Hartley, K. and Huby, M. (1986). Contracting out policy: Theory and evidence, in kay, j., mayer, c. and thompson, d. (eds.). *Privatisation and Regulation: The UK Experience*.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1):153–162.

Holmström, B. (1979). Moral hazard and observability. *The Bell Journal of Economics*, 10:74–91.

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

Jensen, M. and Meckling, W. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3:305–360.

Jensen, P. H. and Stonecash, R. E. (2005). Incentives and the efficiency of public sector-outsourcing contracts. *Journal of Economic Surveys*, 19(5):767–787.

Kvaloy, O. and Olsen, T. (2009). Endogenous verifiability and relational contracting. *American Economic Review*, 99(5):2193–2208.

Kvaloy, O. and Olsen, T. E. (2010). Incentive provision when contracting is costly. *Working Paper*.

Milne, R. G. and Wright, R. E. (2004). Competition and costs: Evidence from competitive tendering in the scottish national health service. *Scottish Journal of Political Economy*, 51(1):1–23.

Savas, E. S. (1977). An empirical study of competition in municipal service delivery. *Public Administration Review*, 37(6):717–724.

Spagnolo, G. (2012). Reputation, competition, and entry in procurement. *International Journal of Industrial Organization*, 30(3):291–296.

Townsend, R. M. (1979). Optimal contracts and competitive markets with costly state verification. *Journal of Economic Theory*, 21(2):265–293.

Wooldridge, J. M. (2001). *Econometric Analysis of Cross Section and Panel Data*, volume 1. The MIT Press.

7 Appendix

Table 1: Descriptive statistics

Variable	Description	Mean	Std. dev.	Min	Max	N
$Quality_{ijt}$	Level of quality supplied by firm$i$ on lot$j$ at time$t$(from 0 to 100)	88.91	6.86	42.5	100	1382
$Observe_{ijt}$	Takes the value 1 if the $Quality_{ijt}$ was measured, 0 otherwise	0.61	0.49	0	1	2248
$NewDesign_{jt}$	Takes the value 1 if the contract for lot j is awarded after April 2010, 0 otherwise	0.56	0.50	0	1	2248
$Decision_t$	Takes the value 1 after the decision of the administrative court of Paris, *i.e.* after May 2009, 0 otherwise	0.78	0.42	0	1	2248
$Penalties_{ijt-1}$*	Overall value of penalties paid by firm i for lot$j$ at time *t-1* (in euros)	563.73	1 531.12	0	13 790	2195
$ControlFreq_{ijt-1}$*	Number of times the buyer controlled the quality supplied by firm$i$ on lot j at time *t-1* / Maximal number of times it could have done it	0.54	0.28	0	1	2195
$Price_{ijt}$	Winning bid of firm i at time$t$ for the fixed part of lot j, divided by the number of accommodations and by the number of months (in euros). This a unit price per month, per accommodation	13.48	3.57	8.94	31.91	102
$NbOffers_{jt}$	Number of offers for lot j at time t	6.23	2.90	1	14	102
$NbAccommodations_{jt}$	Number of accommodations for lot j at time t	1 846.167	593.76	544	3 066	102

* Variables built thanks to information related to periods from January 2008 through June 2008

Figure 1: Quality evolution over time

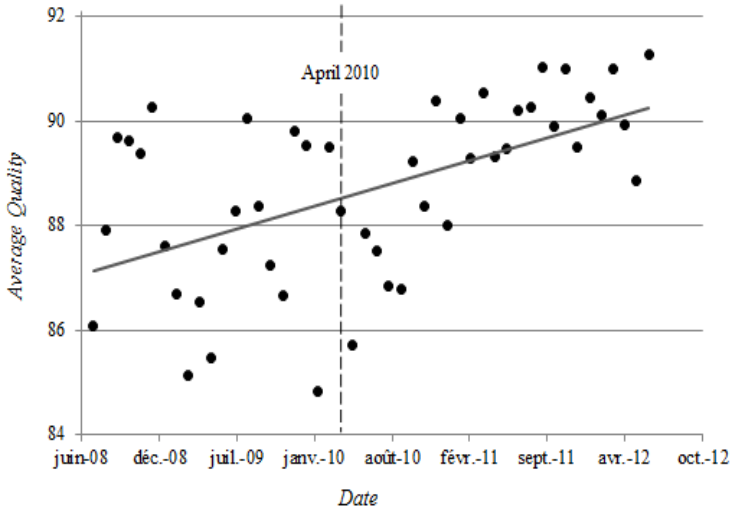

Figure 2: Number of offers and prices evolution over time

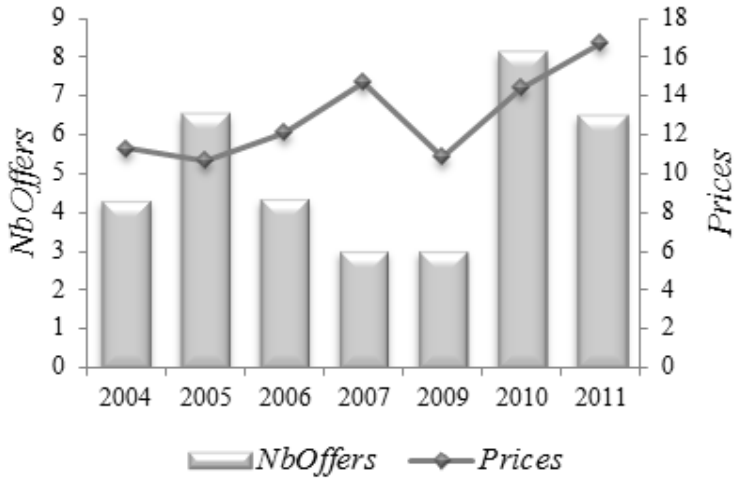

Figure 3: Timing of the events

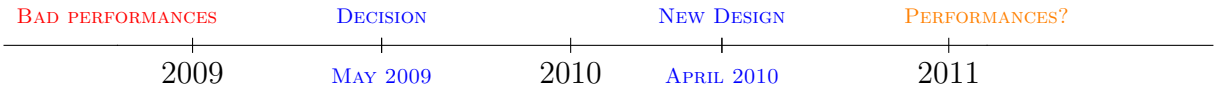

Table 2: Old *versus* new contract design

	Panel A Old contract	Panel B New contract
Tasks descriptions and contractual requirements	62 tasks and 3 levels of frequency (daily, weekly, monthly)	118 tasks and 6 levels of frequency (daily, weekly, monthly, quarterly, semi-annual, annual)
Evaluation for quality	Unchanged	
Performance obligations	Unchanged	
Mandatory documents	Unchanged	
Controls	Contradictory controls (once per month / at the discretion of the public buyer)	Contradictory controls (once per month / at the discretion of the public buyer) + **Unplanned and not contradictory controls** (at the discretion of the public buyer)
Penalties	60 euros if quality score < 80 / 300 euros if the 2nd following quality score is still < 80	**3 %** of the price (fixed part) if quality score < 80 / **6 %** of the price (fixed part) if the 2nd following quality score is still < 80 + **new penalties** for late delivery of mandatory documents

Table 3: Penalties per year

Year	2008	2009	2010	2011	2012**
Penalties*	21646,5	12347,4	37297,2	60641,5	20800,0

* in Euros ; ** until June

Table 4: Sample comparisons

Variable	Panel A	Panel B	t-test*
$Quality_{ijt}$	87.75	89.90	0.0000
$Observe_{ijt}$	0.64	0.60	0.0526
$Penalties_{ijt-1}$	495.64	620.59	0.0569
$ControlFreq_{ijt-1}$	0.64	0.45	0.0000
$Price_{ijt}$	12.31	14.61	0.0009
$NbOffers_{jt}$	4.36	8.02	0.0000
$NbAccommodations_{jt}$	1 839.82	1 852.269	0.9163

* P-value of the difference between means

Table 5: Matrix of correlations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) $Quality_{ijt}$	1								
(2) $Observe_{ijt}$	0	1							
(3) $NewDesign_{jt}$	0.157	-0.0409	1						
(4) $Decision_t$	0.0902	-0.0201	0.601	1					
(5) $Penalties_{ijt-1}$	0.0211	-0.0430	0.0407	0.0934	1				
(6) $ControlFreq_{ijt-1}$	0.0328	0.376	-0.334	-0.188	-0.00336	1			
(7) $Price_{ijt}$ t	0.154	-0.0301	0.299	0.173	0.0809	-0.125	1		
(8) $NbOffers_{jt}$	0.0993	-0.0512	0.629	0.378	-0.0856	-0.272	-0.152	1	
(9) $NbAccommodations_{jt}$	-0.0420	0.0131	0.0222	-0.0202	0.0267	0.000128	-0.0999	-0.0450	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: How to implement quality?

	Model 1 $Quality_{ijt}$	Model 2 $Quality_{ijt}$	Model 3 $Quality_{ijt}$	Model 4 $Quality_{ijt}$
$NewDesign_{jt}$	2.526***	2.006***	2.104***	1.433*
	(0.605)	(0.721)	(0.709)	(0.738)
$ControlFreq_{ijt-1}$			1.529+	-4.283***
			(0.959)	(1.494)
$NewDesign_{jt} * \hat{ControlFreq}_{ijt-1}$				8.601***
				(1.813)
$Penalties_{ijt-1}$			0.000	-0.000+
			(0.000)	(0.000)
$NewDesign_{jt} * \hat{Penalties}_{ijt-1}$				0.001***
				(0.000)
$NbOffers_{jt}$	0.063	0.018	0.088	0.062
	(0.159)	(0.196)	(0.199)	(0.196)
$NbAccommodations_{jt}$	-0.002***	-0.001	-0.001	-0.002+
	(0.001)	(0.001)	(0.001)	(0.001)
$Constant_{ijt}$	91.332***	85.040***	83.855***	88.308***
	(1.505)	(2.288)	(2.474)	(2.696)
Lot	Yes	Yes	Yes	Yes
Firm		Yes	Yes	Yes
N	1359	1359	1359	1359
$Adj - R2$	0.21	0.24	0.24	0.25

+ $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; robust standard errors in parentheses

Table 7: Dealing with sample selection

	Model 5 $Quality_{ijt}$	Model 6 $Quality_{ijt}$	Model 7 $Quality_{ijt}$	Model 8 $Quality_{ijt}$
$NewDesign_{jt}$	2.187***	1.655**	1.685**	1.184+
	(0.611)	(0.688)	(0.713)	(0.734)
$ControlFreq_{ijt-1}$			0.540	-5.192***
			(1.061)	(1.643)
$NewDesign_{jt} * \hat{ControlFreq}_{ijt-1}$				9.238***
				(1.818)
$Penalties_{ijt-1}$			0.000	-0.000
			(0.000)	(0.000)
$NewDesign_{jt} * \hat{Penalties}_{ijt-1}$				0.001***
				(0.000)
$NbOffers_{jt}$	0.128	0.074	0.116	0.076
	(0.133)	(0.161)	(0.164)	(0.163)
$Nb\_Accommodations_{jt}$	-0.002***	-0.001	-0.001	-0.002+
	(0.001)	(0.001)	(0.001)	(0.001)
$Constant_{ijt}$	88.576***	87.731***	86.836***	91.324***
	(1.387)	(2.333)	(2.603)	(2.802)
	$Observe_{ijt}$	$Observe_{ijt}$	$Observe_{ijt}$	$Observe_{ijt}$
$ControlOthers_{ijt-1}$	2.616***	2.681***	2.481***	2.768***
	(0.162)	(0.165)	(0.174)	(0.180)
$NewDesign_{jt}$	0.374***	0.460***	0.526***	0.651***
	(0.110)	(0.124)	(0.127)	(0.130)
$ControlFreq_{ijt-1}$			0.656***	2.303***
			(0.170)	(0.277)
$NewDesign_{jt} * \hat{ControlFreq}_{ijt-1}$				-2.550***
				(0.335)
$Penalties_{ijt-1}$			-0.000	-0.000
			(0.000)	(0.000)
$NewDesign_{jt} * \hat{Penalties}_{ijt-1}$				-0.000
				(0.000)
$NbOffers_{jt}$	-0.088***	-0.093***	-0.085***	-0.107***
	(0.023)	(0.029)	(0.030)	(0.031)
$Nb\_Accommodations_{jt}$	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
$Constant_{ijt}$	-0.439	-0.506	-0.895*	-2.049***
	(0.312)	(0.488)	(0.501)	(0.537)
Lambda	-1.923**	-1.873**	-1.726**	-0.861
	(0.775)	(0.746)	(0.815)	(0.753)
Lot	Yes	Yes	Yes	Yes
Firm		Yes	Yes	Yes
N	1359	1359	1359	1359

+ $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; robust standard errors in parentheses

Table 8: Testing for the exogeneity of the change in contract design

	Model 9 $Quality_{ijt}$	Model 10 $Quality_{ijt}$	Model 11 $Quality_{ijt}$	Model 12 $Quality_{ijt}$
$Decision_t$	0.310	0.297	0.233	0.591
	(0.499)	(0.497)	(0.500)	(0.499)
$NewDesign_{jt}$	2.014***	1.497**	1.562**	0.877
	(0.670)	(0.738)	(0.759)	(0.777)
$ControlFreq_{ijt-1}$			0.529	-5.319***
			(1.062)	(1.645)
$NewDesign_{jt} * \hat{ControlFreq}_{ijt-1}$				9.341***
				(1.820)
$Penalties_{ijt-1}$			0.000	-0.000+
			(0.000)	(0.000)
$NewDesign_{jt} * \hat{Penalties}_{ijt-1}$				0.001***
				(0.000)
$NbOffers_{jt}$	0.129	0.075	0.115	0.070
	(0.133)	(0.161)	(0.164)	(0.163)
$NbAccommodations_{jt}$	-0.002***	-0.001	-0.001	-0.002+
	(0.001)	(0.001)	(0.001)	(0.001)
$Constant_{ijt}$	88.272***	87.457***	86.655***	90.936***
	$Observe_{ijt}$	$Observe_{ijt}$	$Observe_{ijt}$	$Observe_{ijt}$
$Decision_t$	0.007	-0.011	0.006	0.019
	(0.091)	(0.091)	(0.093)	(0.097)
$ControlOthers_{ijt-1}$	2.616***	2.681***	2.481***	2.769***
	(0.162)	(0.165)	(0.174)	(0.180)
$NewDesign_{jt}$	0.371***	0.466***	0.523***	0.641***
	(0.120)	(0.134)	(0.136)	(0.139)
$ControlFreq_{ijt-1}$			0.656***	2.305***
			(0.170)	(0.277)
$NewDesign_{jt} * \hat{ControlFreq}_{ijt-1}$				-2.552***
				(0.335)
$Penalties_{ijt-1}$			-0.000	-0.000
			(0.000)	(0.000)
$NewDesign_{jt} * \hat{Penalties}_{ijt-1}$				-0.000
				(0.000)
$NbOffers_{jt}$	-0.088***	-0.093***	-0.085***	-0.107***
	(0.023)	(0.029)	(0.030)	(0.031)
$NbAccommodations_{jt}$	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
$Constant_{ijt}$	-0.446	-0.494	-0.901*	-2.067***
	(0.326)	(0.497)	(0.510)	(0.545)
$Lambda$	-1.928**	-1.874**	-1.741**	-0.906
	(0.775)	(0.746)	(0.816)	(0.753)
Lot	Yes	Yes	Yes	Yes
Firm		Yes	Yes	Yes
N	1359	1359	1359	1359

+ $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; robust standard errors in parentheses

Table 9: Prices (bids) and change in contract design

	Model 13 $Price_{ijt}$	Model 14 $Price_{ijt}$	Model 15 $Price_{ijt}$
$NewDesign_{jt}$	2.203	5.579	0.253
	(2.898)	(3.898)	(3.459)
$NbOffers_{jt}$	-0.706***	-0.469**	-0.733***
	(0.181)	(0.198)	(0.190)
$NbAccommodations_{jt}$	-0.001	-0.001	-0.001
	(0.001)	(0.002)	(0.001)
$PriceIndex_t$	0.155	-0.098	0.226
	(0.128)	(0.169)	(0.178)
$Constant_{ijt}$	1.196	26.947	-5.142
	(13.587)	(20.161)	(18.138)
Lot		Yes	
Firm			Yes
N	102	102	102
$Adj-R2$	0.34	0.49	0.38

+ $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; robust standard errors in parentheses