Relational Contracts, Procurement Competition, and Supplier Collusion*

Giacomo Calzolari† Giancarlo Spagnolo‡

This version: February 20th, 2017

Abstract

We study the tension between competitive screening, contract enforcement and supplier collusion where a buyer trades repeatedly with one among several suppliers, moral hazard and adverse selection coexist, and non-contractible tasks are governed by relational contracting. Open competition is optimal when suppliers are few and heterogeneous and non-contractible quality is not too important. Otherwise, the buyer optimally restricts competition to a subset of regular and frequently interacting suppliers. These policies facilitate suppliers' collusion but we show that the buyer may benefit from it enforcing even higher non-contractible tasks. These results shed light on a number of puzzling observations connecting the worldwide "auto parts" cartels, Toyota's "lean" management practices and the persistent trade imbalance between Japan and the US in the 1990s.

JEL Classification Numbers: C73, D86, L14.

Keywords: Relational contracts, Non-contractible quality, Screening, Collusion, Limited enforcement, Lean procurement, Management practices, Supply chains.

*We are grateful for their helpful comments to Pierpaolo Battigalli, Pierre-André Chiappori, Vincenzo Denicolò, Mathias Dewatripont, Nicola Doni, Bob Gibbons, Sergei Izmalkov, Jonathan Levin, Bentley Macleod, Alejandro Manelli, Claudio Mezzetti, Sebastien Mitraille, Antonio Nicolò, Yossi Spiegel, Daniel Traca, Andy Skrzypacz, Jeanine Thal and Paola Valbonesi. We thank seminar participants at numerous workshops, conferences and seminars. Spagnolo is grateful to the Swedish Research Council (Vetenskapradet) for research funding.

†Department of Economics, University of Bologna, Italy and CEPR. E-mail: *giacomo.calzolari@unibo.it*

‡University of Rome 'Tor Vergata', SITE, EIEF & CEPR. E-mail: *giancarlo.spagnolo@uniroma2.it.*

1 Introduction

The success of Toyota in the last half-century is widely attributed to its "lean" management practices based on relational contracting.[1] In procurement, these imply long-term relations with a restricted group of dedicated suppliers, instead of open competitive procurements. Relational contracts have even found their way into the traditionally more rigid public procurement, reducing reliance on open competitive auctions and leaving more scope for flexible, restricted auctions with requests of quotations to a subset of potential suppliers and strong influence of past performance and "the shadow of the future".[2]

Relational contracts allow parties to govern sometimes crucial non-contractible dimensions of the supply relation through self-enforced cooperation. Cooperation incentives are typically stronger the larger the expected future payoffs at stake, so that relational enforcement may conflict with other important needs of a principal, in particular that of having suppliers compete to screen and select the more capable and least costly among them. In this paper we study this potential conflict in a model where a buyer/principal trades recurrently and sequentially with multiple competing suppliers/agents and where non-contractible dimensions are important. We characterize the buyer's optimal stationary relational contract, defined in a broad sense to include equilibrium choices on explicitly contracted features, on non-contractible dimensions, and on the competitive screening policy. We identify situations in which tension does arise between competition and enforcement, deriving the general implications of this trade-off. We show that this tension can be so marked that actually negating competition and procuring with colluding suppliers may become optimal for a buyer in quest of the highest non-contractible performance.

By connecting relational management practices to suppliers cooperation and cartels, our analysis helps to shed light on unexplored features of relational contracting in procurement and on the relationship between these and the large set of recently discovered supplier cartels in the auto-part industry. It also offers an explanation of why US cart-parts suppliers struggled to enter the Japanese market leading to the heated trade disputes of the 1990s.

Outline of the model and results. In our setting, an infinitely-lived buyer/principal procures recurrently a task with a non-contractible dimension from a population of heterogeneous and

[1] Helper and Henderson (2014) offers a brilliant account and many references on this saga.

[2] See for example in the US, the reforms started under Kelman and the 1994 Procurement Streamlining Act has led to an increase use of flexible procedures with limited competition (Kelman, 1990).

privately informed, infinitely lived suppliers/suppliers. Both moral hazard on non-contractible dimensions and adverse selection on suppliers' type/cost are present. When the principal designs the optimal relational contract, he must choose: any explicit part of the contract enforceable by courts, such as contractible payments, contract duration and contractible performance standards; the implicit and self-enforcing incentives designed to govern costly non-contractible performance; and the auction format and screening policy, i.e. how and how often suppliers compete for procurement.

We find that restricting competition to a stable pool of eligible, trusted suppliers under the threat of exclusion and replacement in case of under-performance, as is typical of the "Toyota way", may optimally solve the principal's problem of competition and enforcement. To enforce substantial levels of the non-contractible task, the buyer has to rely on promised large future rents by limiting future competition among those pre-selected suppliers, trading off the cost of lower efficiency and higher costs.

Investigating this first trade-off, we show that restricted competition is more likely to be optimal over open competition where all potential suppliers are invited to compete, (i) the more valuable the non-contractible task is; (ii) the larger the number of potential suppliers in the market; (iii) the smaller suppliers' cost heterogeneity; and (iv) the smaller the common discount factor, interpreted as a lower frequency of interaction/longer contract duration.

These results are consistent with the way Asanuma (1989), Aoki and Dore (1994) and Taylor and Wiggins (1997), among others, have related the "Toyota way" to lean, relational procurement: requests for proposals (auctions) restricted to few – often two – loyal long-term suppliers; small batches inducing high frequency of interaction; and ex post quality evaluation rather than ex ante supply inspection. The comparative statics also seem to fit well with the switch to fewer, trusted suppliers observed in IT supply chains after globalization drastically increased the number of potential suppliers (Bakos and Brynjolfsson, 1993), while the exclusionary features of the relational contracts we characterize may explain the difficulties faced by American auto parts producers in accessing Japanese markets in the 1990s even after formal trade barriers were removed (see Cooper, 2014, for a retrospective).

Although restriction to a smaller pool of frequently interacting suppliers facilitates cooperation between a buyer and his suppliers, at the same time it also facilitates suppliers' collusion against the buyer. We therefore proceed to illustrate a second and rather general trade-off between relational enforcement and supplier collusion which disappointingly seems to denote limits to the possible remedies for non-contractible tasks. These results are suggestive of why, just at the time several car manufacturers have converged to Toyota-style relational procurement, a large number of overlapping cartels among automotive parts suppliers – most of which Toyota suppliers – have been recently uncovered by antitrust authorities (more details in Section 2).[3] This set of overlapping cartels, still in the middle of being investigated, has already led to a record amount of fines and jail sentences in what is emerging as the biggest antitrust investigation ever.

However, our analysis clarifies that collusion itself may interact directly with suppliers' incentives for non-contractible performance. By increasing the price of the contract, collusion increases suppliers' gains from future trade with the principal which in turn represent the opportunity cost of being excluded for poor non-contractible performance. We show that there are circumstances in which the trade-off between non-contractible performance and collusion is only apparent. When contractible dimensions are very important and negotiation with a single supplier is not allowed, or it is too risky for the danger of unexpected disruption in the procurement flow, it may actually be optimal for the principal to induce few trusted suppliers to fix the supply price. These results help explain a number of puzzling features of the auto parts cartels which would be otherwise difficult to explain with alternative models.

We generalize the analysis allowing parties to also exchange contractible monetary ex ante transfers – like wages and participation fees – as well as non-contractible ones, like discretionary performance bonds and bonuses. Discretionary transfers may in principle allow a buyer to govern non-contractible dimensions limiting the role of information rents to suppliers, thus pushing towards a more extensive use of open competition. However, the principal's temptation to renege on a promised bonus (or to withhold a bond), together with the possibility to exclude the betrayed performing supplier, limits the use of such discretionary transfers. Taking this incentive problem into account, we find that when suppliers compete and discretionary transfers are available, the principal optimally chooses bonuses and open competition when there are few available suppliers in the market and non-contractible dimensions are not too important. Still, restricted competition

[3]The Economist, "Cartelbusting - Boring can Still be Bad,", "Just One More Fix," March 29th 2014; "No truck with cartels" Jun 4th 2016; "Massive price-fixing among auto-parts manufacturers hurt U.S. car buyers," The Columbus Dispatch, March 22nd, 2015.

The first version of the present paper circulated several years before the car part cartel was discovered. To avoid being sued for providing criminal inspiration to practitioners, we notice that most of those conspiracies started a few years before that and none of the authors ever worked for Toyota, other car manufacturers and any of their suppliers.

is optimal when the non-contractible task is important and there are many potential suppliers. Hence, the enforcement-collusion trade off and the desirability to procure with colluding suppliers in some situations are reaffirmed with a very rich set of contractual instruments.

Results are also consistent with recent evidence on the use of restricted auctions in public procurement.[4] They also shed light on the *Kansei-Dango* practice in Japan whereby, according to the Japan Fair Trade Commission, public buyers traditionally coordinated bid-rigging among suppliers – even in the absence of corruption – to avoid price competition potentially disrupts the quality of procured goods and services;[5] and on the analogous and more general tendency of public procurers to soften price competition, allowing for consortia and restricted procedures when procurement is complex and non-contractible aspects crucial.[6]

Finally, in light of Toyota's investment in creating and maintaining the Bluegrass Automotive Manufacturing Association (BAMA) and the associated regular information-sharing meetings among buyer and suppliers (Milgrom and Roberts, 1993), we also consider the case of multilateral relational contracting. The principal may, at some cost, create an information-sharing device (e.g. like the BAMA) that ensures that the level of non-contractible quality provided by a supplier and the possible problems with performance bonuses become known to all suppliers in the pool.[7] As in Levin (2002), this allows the principal and the associated suppliers to establish *multilateral relational contracts* where, if the principal reneges on a promise, all suppliers in the pool stop cooperating. Multilateral relational contracts improve the principal's ability to commit not to renege on a promised bonus, ensuring that even higher non-contractible quality is sustainable. Being communication and information-sharing, as in BAMA, well known factors that facilitate suppliers' collusion, we also explore the consequences of collusion for optimal multilateral relational procurement.

Related literature. Our paper is related to several strands of literature. It contributes to the large literature on relational contracts that stems from the pioneering contributions of MacLeod

[4]See Section 8, Coviello et al. (2017) and references therein.

[5]General Secretariat, Japan Fair Trade Commission, 2015, "Towards the Prevention of Bid-Rigging", p. 68.

[6]Of course, relational contract may end up involving corruption, besides collusion, but the recent work of Troya-Martinez and Lewis (2016) suggests this may further increase efficiency

[7]Among other things, BAMA "...provides expanded opportunities for direct communications with Toyota executives and access to Toyota training programs, while giving members the opportunity to share best practices and collectively raise issues and communicate problems affecting the way we do business." (from http://www.bama-group.org/, October 7th, 2013).

and Malcomson (1989) and Baker, Gibbons and Murphy (1994). The work closest to our own is probably that of Levin (2003), who elegantly characterizes the optimal relational contract with moral hazard and adverse selection between a buyer and a single seller. A major difference is that we focus on competition among several suppliers and we are thus concerned with the trade-off between competitive screening and the incentives to deliver non-contractible performance. Our work is also close to MacLeod and Malcomson (1998), which posits relational contracts between a number of principals and a larger number of competing agents, although without adverse selection and screening. Taylor and Wiggins (1997) provided a first model of the Japanese-style relational procurement focusing on inspection costs and the frequency of interaction between a buyer and a single supplier.

So far, the literature has paid limited attention to the case of a relational contract between one principal and several privately-informed suppliers who compete recurrently to be selected for a procurement contract. Levin (2002) studies team production in a relational environment where agents do not compete. A similar environment is that in Rayo (2007) who studies relational contracts with multiple agents, endogenously deriving the organizational structure, but again in a framework with team production and no competition.[8] Board (2011) obtains a result analogous to ours on the optimality of limiting the number of trading partners, but in a very different model where the focus is on the dynamics of the relational contract and where there is no need to competitively screen suppliers while contractible monetary transfers (e.g. wages, prices) and performance bonuses (or bonds) are not admitted. In the same vein, a recent paper by Andrews and Barron (2016) focuses on the dynamics of contract allocation among multiple agents as a device to make the principal's promises of discretionary bonuses credible, but does not discuss competition nor collusion among suppliers. Our paper is complementary to these two in the sense that, focusing on stationary contracts, we can characterize the roles of competition and collusion among suppliers.

Our paper is also related to the literature on the optimality of using open, competitive auctions. When contracts are complete and their enforcement costless, classic results on the optimality of open auctions apply (e.g. Bulow and Klemperer, 2009). When contracts are highly incomplete and costly to enforce, i.e. when non-contractible quality is important, open auctions may perform

[8]Mukherjee and Vasconcelos (2011) compare individual task responsibility with team responsibility, and show that teams soften the multi-tasking problem but weaken relational contracting. They do not consider competition nor collusion among agents.

poorly in terms of purely economic outcomes. Focusing explicitly on construction procurement, Spulber (1990) showed how the interplay of incomplete contracting and auctions may intensify problems of moral hazard and ex post opportunism, leading to rather poor outcomes. Manelli and Vincent (1996) reach an even more extreme conclusion, showing that when the crucial dimension on which gains from trade are concentrated is not contractible, open auctions that induce bidders to compete on contractible dimensions (e.g. price) are the worst among conceivable procurement mechanisms, as they maximize the damages from adverse selection.

In a dynamic framework, auctions with a choice of participant depending on past performance may allow the buyer to take into account reputational forces and establish long-term relationships that may improve performance (Kim 1988, Doni 2006). Our paper shows how discretion to restrict competition may allow a buyer to sustain relational contracts that enforce non-contractible quality, but also that this solution inevitably facilitates collusion among suppliers, which can have positive or negative effects on procurement outcomes.

Our analysis is also relevant to the literature on reputation and markets. Stiglitz (1989) raised the question of how reputation could be compatible with perfect competition, which should eliminate all future supracompetitive gains (see also Kranton 2003, suggesting that professional associations may have a role in limiting competition, and also Bar-Isaac 2005). Hörner (2002) offered an elegant answer to Stiglitz's question. In his model with heterogeneous consumers and firms, adverse selection (time-persistent costs) and moral hazard (quality of goods), high prices signal high quality and make competition compatible with reputational forces. In our environment signalling would not work because of time-varying costs and because prices are set in a competitive auction (or, in negotiations, they are determined by the buyer).

Finally, the fact that our model with limited enforcement offers an explanation for the trade frictions between the US and Japan in the nineties links our paper to the strand of literature that, starting with Antras (2003) and Antras and Helpman (2004), introduces contractual frictions typical of firms supply chains in an international trade context.

The rest of the paper is organized as follows. Section 2 illustrates the puzzling auto-parts cartel that has been recently unveiled. Section 3 describes our model. Section 4 studies the optimality of restricted competition and Section 5 discusses the risk and consequences of suppliers' collusion. Section 6 investigates the possibility for the buyer to offer ex-ante transfers and discretionary payments. In Section 7 we discuss extensions and some robustness analysis. Section 8 concludes.

2 The auto parts cartel

Over the last couple of years, competition authorities from 15 separate jurisdictions, including that in Europe, the US, and Asia, have undertaken over 30 investigations into anticompetitive conduct in the automotive parts sector. The United States Department of Justice (DoJ) has described its investigation as "the largest criminal investigation the Antitrust Division has ever pursued," both in terms of its scope and the potential volume of commerce affected by the alleged illegal conduct (Goldfein and Keyte, 2014). Apparently, it all started in February 2010 when Sumitomo, a Japanese supplier producing wire harnesses, applied for a marker under the Notice on immunity from fines in cartel cases, the so called Leniency Note. According to Brent Snyder, head of the criminal enforcement efforts at the Department of Justice, the DoJ began raiding the companies in 2010 following the leniency application. Companies used code names, met in remote locations to fix the prices of starter motors, seat belts, radiators and more, and followed up with each other "to make sure the collusive agreements were being adhered to," the DoJ alleges. Twenty-six firms, many of them Japanese, have already pleaded guilty and agreed to pay a total $2 billion in fines. By 2014 two dozen people had been charged, and much more is under way. Cases brought involved about 30 parts, but the DoJ believes the prices of 100-150 part may have been manipulated. Cartels are estimated to have started at the beginning of 2000, and to have lasted unusually long periods ranging from 5 to 14 years, which may even be optimistic estimates according to the discussions of the US-Japanese "trade war" in the 1990s. The focus of the investigations has mainly been on Japanese automakers and suppliers, and many casual observers seem to understand that in Japan there is a prevalent cultural history and tradition of doing business in ways that may be conducive to collusion.

The auto parts cartels present some puzzling features when compared to other cartel stories. A first puzzling aspect is that automakers are very large, few in number, traditionally tough, sophisticated, and very well informed buyers. These are all textbook conditions that make life very hard for secret cartels among suppliers and a long duration of such cartels very unlikely. The second is that these cartels seem to have had unusually long duration. A third puzzling feature is that many of the suppliers at the core of these cartels were almost completely controlled by their buyers, and in particular by Toyota. A fourth puzzling feature is that there seems to be no sign of these cartels harming the performance of Japanese automakers. To check for this, we downloaded all the decisions available from the homepage of the Antitrust Division of the

Department of Justice relative to 21 auto part suppliers. It is interesting to note that all the 21 firms happened to be regular suppliers of Toyota. Only 12 of them have been suppliers for Honda and Nissan, six have been suppliers for Ford, and only one seems to have been a supplier for most other auto manufacturers. Figure 1 plots the recent profitability of large car manufacturers, including Toyota. While the alleged auto part cartels have been active, Toyota seem to have had the best performance of the market. This suggests that these cartels did not significantly hinder the appeal of Toyota's cars to consumers and the company's profitability, although we do not have the counterfactual of what would have happened in the absence of such cartels. A fifth puzzling observation is that among car manufacturers, only Ford sued some of the convicted suppliers for damages.

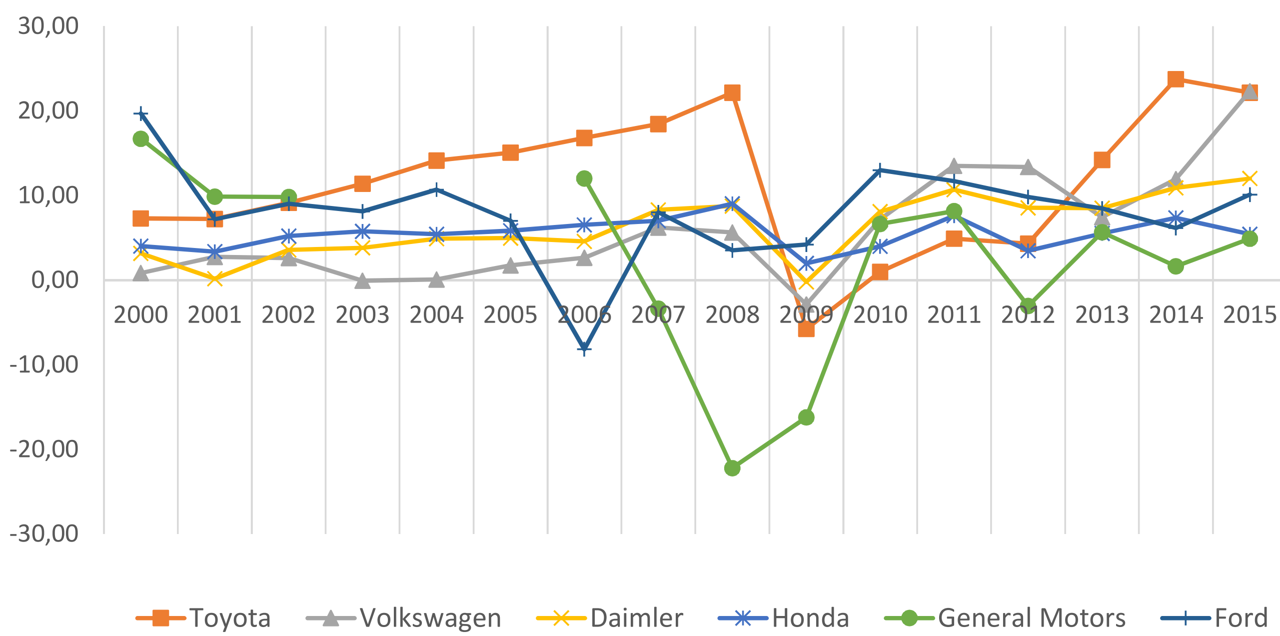

Figure 1: Operating income of large automakers (billions of US$)

In light of these puzzles, inevitable questions emerge. Were these cartels really harming automakers? Were they harming consumers? Were automakers really unaware of these concerted practices? Or were they part of the "managed competition" style of the "Toyota-way", which inspired the huge management literature (and consulting cottage industry) on "lean production"?

Providing consistent answers to all these questions seems rather challenging. We believe that our model manages to provide answers that differ for Toyota than for automakers that are less keen on lean management practices. Our results suggest that these collusive practices were likely to be known, tolerated, and even encouraged by Toyota as an instrument to ensure quality of auto parts supply. They also suggest that these vertical arrangements, while exclusionary for outsider (e.g. US) auto part producers, may have benefitted (at least some) consumers, who have continued to favor the high-quality Toyota cars. We will discuss potential negative spillovers in Section 8.[9]

3 The model

Environment. At any period $t = 0, 1, 2, ...$ a buyer (principal), needs a task to be performed by one among $N > 1$ potential suppliers (agents). The principal's per-period value of the task $v(q_t)$ is increasing and concave in a costly decision q_t (≥ 0) taken by the supplier of period t. The per-period cost for supplier i is $\theta_{it} + \psi(q_t)$ with $\psi(q_t)$ increasing, differentiable, strictly convex and $\psi(0) = \psi'(0) = 0$. The value of trade in period t with supplier i is $s(q_t) - \theta_{it}$ where $s(q_t) \equiv v(q_t) - \psi(q_t)$. With an adequate measure for q_t and appropriately scaling ψ, we set $v(q_t) = vq_t + v_0$ with $v > 0$ and $v_0 \geq 0$. The time horizon is infinite, and all players are risk-neutral and have a constant and common discount factor $\delta \leq 1$.

Although q_t is observable to the principal and the supplier, it is not verifiable to third parties, which makes it non-contractible. For example, q_t may be the costly effort (or specific investment) provided by an expert or an employee, or a quality feature of the procured service that cannot be fully specified. We will refer to q_t as non-contractible "quality". In Section 8 we will partially relax the assumption on non-verifiability by allowing it to be observed by some of the suppliers.

In this set up, supply can be seen as an indivisible multi-tasking activity which contemplates a contractible decision taken by the supplier at cost θ_{it} that is worth v_0 to the principal and generates value $v_0 - \theta_{it}$, and a non-contractible decision q_t. To avoid uninteresting cases, we assume $v_0 > \theta_{it}$ for any θ_{it}, so that the principal never wants to discontinue supply.

The suppliers' outside option is $\underline{u} = 0$ and the principal's $\underline{v} \geq 0$.

[9] Although it is difficult to imagine cartels larger and broader in scope than the those of auto parts, it is certainly not the only case where cartels were following up an attempt by buyers to stop a competitive race to the bottom on non-contractible dimensions. For example, in the UK Steel Tank Cartel case (Stephan, 2016), a cartel was initiated after competitors were asked to meet up by a buyers' industry body to discuss minimum safety standards following many complaints and a period of price wars. The defendants have been arguing forcefully that the main raison d'être of the cartel was safeguarding safety standards on request from the buyers.

As in Levin (2003), we assume that $\theta_{it} \in \Theta \equiv [\underline{\theta}, \overline{\theta}]$ is independently and identically distributed with density $f(\theta_{it})(> 0)$ and $\theta_e \equiv E(\theta)$ and all this is common knowledge. In Section 8 we will discuss the case of persistent costs. In any period, the principal is uninformed about θ_{it} and, to simplify exposition, we assume suppliers are instead fully informed, as knowing each other well is a common circumstance in procurement or among colleagues in a firm. This assumption, immaterial for the trade-offs underlying our results, will be further discussed in Section 8.

Since the principal is uninformed about θ_{it}, in any t he uses a competitive screening device, an auction with rules A_t, that awards an explicit contract requiring the winning supplier to provide the specified supply for one period.[10] To fix ideas, we refer explicitly to first price auctions, but other standard auction formats would do for our analysis as well.[11] The auction rules contemplate the price paid to the auction winner for supplying, i.e. the lowest winning bid b_t' and, in some cases, also a reservation price r. In addition to these "interim" verifiable monetary transfers, in Section 6 we will allow for transfers exchanged ex ante, i.e. before suppliers learn their cost and compete (like wages or participation fees, denoted by w_t), and for ex post discretionary transfers paid after the execution of the contract (such as bonuses and bonds, B_t).

Explicit and Relational Contracts. Although quality is not contractible, the principal and the suppliers may still profit from ongoing interactions and reach an implicit agreement, or "*relational contract*", on how they are going to behave in the future, on and off equilibrium. A relational contract is self-enforcing if it describes a perfect public equilibrium of the repeated game (Fudenberg, Levine and Maskin, 1994). In the following we use simply the "*contract*" to refer to the explicit court-enforced contract awarded at the auction stage and regulating the verifiable dimensions of the exchange, i.e. the price of supply b_t' established at the auction, the basic task/provision with value v_0 and the auction rules A_t itself. We thus distinguish the contract from the implicit self-enforcing part of the relational contract.

We will first focus on "bilateral relational contracts", denoted by C_i and defined as complete and independent contingent plans of action for the principal and supplier i as in MacLeod and Malcomson (1989). Upon a deviation at date t, either by the date-t supplier or by the principal,

[10] Although not explicitly modelled here, the principal could offer contracts that last for more than one period (or renew current ones). This is captured with an increase of the discount factor δ. Calzolari and Spagnolo (2009) model contract duration explicitly with no substantial differences.

[11] Agents may compete announcing a quality and an associated price, with the principal ranking offers by a scoring rule. This is irrelevant in our setup because of the same cost of quality for all agents.

the bilateral relational contract C_i prescribes how parties will behave in the future on and off the equilibrium path. In the present set up, when a deviation occurs in the relational contract C_i, none of the other players knows that the deviation took place, nor can they ascertain the identity of the deviator if they become aware that a deviation took place. This assumption will be relaxed in Section 7 when we will let the suppliers who participated in the auction share some information, extending the analysis to multilateral interdependent relational contracts, as in Levin (2002).

Screening modes. Since the principal may gain by restricting participation, he forms *a pool* of $n_t \leq N$ suppliers invited to compete in the auction at date t, with a process discussed next. We accordingly define the following.

Definition 1 (Screening modes) *With* open competition, *the principal allows all suppliers to participate in the auction and sets $n_t = N$. With* restricted competition, *the principal restricts participation to $n_t < N$ invited suppliers, which also includes the limit case of (bilateral)* negotiation *when $n_t = 1$.*

The number of invited suppliers, whether it is N or smaller, is observable by those invited to compete but not verifiable, thus being part of the relational contract.[12]

Summarizing, the time line of decisions in each period t is as follows:

Time t				Time t+1
Principal sets A_t and invites n_t suppliers	Invited suppliers bid, the winning bid b_t' is paid to the supplier	Winning supplier chooses q_t	Principal observes q_t	... ...

3.1 Benchmarks

Contractible quality. Temporarily assume that quality is verifiable and can be contracted upon. Since the environment is stationary, optimal procurement is also stationary.[13] The principal optimally sets q_{FB} that maximizes the surplus $s(q)$, and suppliers compete for the contract that

[12] Although not always realistic, assuming that n_t is explicitly contractible would strengthen our results. Assuming that agents observe n_t only after bidding would not affect our results.

[13] The proof of this standard result is, for example, in Laffont and Tirole (1993, pages 103-105), for a very similar environment.

contemplates procurement of such quality. With $n > 1$ competing suppliers and given a realization of the costs $\theta_t = (\theta_{1t}, ..., \theta_{nt})$, the most efficient supplier wins bidding a price b' which corresponds to the total cost of procuring q_{FB} for the second-most efficient supplier. Hence, at any auction each of the n invited suppliers expects to earn a profit $\beta(n)\pi(n)$ where $\pi(n)$ is the supplier's expected *informational rent* $\pi(n) \equiv \theta(n) - \theta'(n)$, with $\theta(n)$ and $\theta'(n)$ being respectively the expected cost of the second-most efficient supplier and of the most efficient supplier out of n competitors, and $\beta(n) = \Pr[\theta_{it} = \theta'(n)] = 1/n$ is the probability of being the most efficient one.[14] Since the expected total cost $\theta(n) + \psi(q_{FB})$ of procuring is decreasing in n, the principal prefers open competition and invites all N potential suppliers to bid. Hence, with contractible quality the procurement contract C_{FB} is efficient with $q = q_{FB}$ and $n = N$ at any t and, after substituting for the winning bids, the associated principal's (expected) payoff is

$$V_{FB} = [s(q_{FB}) - \theta(N)]\frac{1}{1-\delta}.$$

If the principal was also fully informed on suppliers' costs, she could directly contract with the most efficient supplier in any period with a take-it-or-leave-it offer. The only difference would be that the principal leaves no rent to the supplier and faces a lower procurement cost equal to $\theta'(N) + \psi(q_{FB})$.

"Zero-quality" equilibria. Back to non-contractible quality, and as usual in relational contracting, there always exist equilibria in which the principal expects $q_t = 0$ and suppliers indeed procure zero quality. Also in this case, the principal would not gain from limiting competition and thus optimally sets $n_t = N$ in any period with an associated payoff

$$V_0^* = [v_0 - \theta(N)]\frac{1}{1-\delta}.$$

4 On the optimality of restricted competition

We now study under what conditions equilibria emerge that allow the principal to implement strictly positive non-contractible quality. We will focus on stationary relational contracts that specify the same auction rules A, non-contractible quality q and the pool of invited suppliers n at any period on the equilibrium path. The stationary equilibria we characterize can be seen

[14] To simplify notation, but without loss of generality, we will proceed as if the event that two or more agents have exactly the same cost has zero measure.

as the long-run steady-state equilibrium paths that follow possibly non-stationary initial phases, like those studied in Board (2011) where invited agents are selected sequentially in a pool that converges to n.[15]

When a relational contract prescribes a strictly positive non-contractible quality q, suppliers may be tempted to cheat and save on quality costs with two types of deviation. First, the supplier that has won an auction may decide to deliver zero quality on the current contract. Second, anticipating that he will cheat on quality in case of being awarded the contract, a supplier may bid aggressively also at a price lower than his costs, and win the contract even when not the most efficient.

The principal can control and obtain (positive) quality by rewarding suppliers with the expected rent from participating in future auctions and punishing deviating suppliers by excluding them from such a rent. In particular, the principal may establish with each supplier i a bilateral relational contract C_i^n that contemplates the following: (i) there are no more than n suppliers invited to compete, (ii) the winning supplier must provide non-contractible quality $q > 0$, (iii) the principal will continue inviting him to future auctions when the quality he actually provided is at least q, and (iv) if instead the supplier deviates ("cheats") by providing quality lower than q, then he will be excluded from the pool for at least one period and possibly replaced by a newly invited supplier if $n < N$.[16]

We *temporarily* assume that the principal can commit to the actions prescribed in the set C^n of n bilateral relational contracts C_i^n.

Lemma 1 (Agents' incentives) *Assume the principal can commit. A relational contract induces the n invited suppliers to participate, to bid with the intention to deliver q, and then to deliver q upon winning an auction, if and only if*

$$\beta(n)\pi(n)\frac{\delta}{1-\delta} \geq \psi(q). \tag{1}$$

Condition (1) guarantees that no supplier has an incentive to deviate by providing a quality lower than q. This is the case both for the most efficient supplier at any t and for any other

[15] In Board's preliminary non-stationary phase, insiders are treated differently than other agents, a discriminatory treatment that is unfeasible in most public or regulated procurement, where accountability requires equal treatment of all bids. A buyer facing these restrictions may optimally "jump" to our stationary steady state directly.

[16] We do not consider milder punishments such as partial handicapping instead of exclusion because they are dominated in cases of interest, i.e. when the buyer wants to procure significant non-contractible quality and because suboptimal with bonuses.

supplier who may conceive of outbidding the most efficient one and subsequently cheat on quality. Indeed, the condition guarantees that what any supplier loses in the future by being excluded (the l.h.s. of (1)) is higher that what he can save by supplying but not providing the quality today (the cost in the r.h.s.). Suppliers anticipate that if they cheat they sacrifice the present value of any expected future profits, since the first supplier who deviates is excluded for at least one period but in fact remains out of the pool forever. This establishes sufficiency. Necessity simply follows from the strategy profile being an optimal penal code in the sense of Abreu (1988).

From inequality (1), any q is *enforceable*, i.e. it is incentive compatible and guarantees suppliers' participation, if and only if $q \leq q(n)$ where

$$q(n) = \psi^{-1}[\beta(n)\pi(n)\frac{\delta}{1-\delta}] \tag{2}$$

is defined as the *maximum enforceable quality* that the principal can procure for any n. Note that since both $\beta(n)$ and $\pi(n)$ are decreasing in n, the maximal enforceable quality $q(n)$ decreases with the size n of the pool of invited suppliers.

The principal's program with commitment can then be written as

$$(\mathcal{P}_c) \qquad \max_{N \geq n,\ q(n) \geq q \geq 0} \quad V(n,q)$$

where

$$V(n,q) = [s(q) - \theta(n)]\frac{1}{1-\delta}. \tag{3}$$

We denote denote with n^* and q^* the solution of this principal's program with commitment $(\mathcal{P}_c)$.

Consider now *the principal's incentives*, thus abandoning the assumption of commitment. Since in our environment the Nash equilibrium of a first price auction leads to the best possible outcome for the principal (for given q and n), there is no way she can gain by modifying the rules of the auction A. Similarly, the principal cannot gain by replacing a supplier who did not cheat in the pool with a new one if available (i.e. when $n < N$), because suppliers are ex ante identical. If $n < N$, the principal could gain by increasing n so as to improve competitive screening and reduce the expected price paid and cost $\theta(n)$. However, any such deviation would be observed and met in the same period by the suppliers who could then begin supplying nil quality. Finally, the principal may prefer not to exclude a supplier who procured a quality lower than q, differently from what is prescribed in the relational contract C^n. We shall say that a relational contract C^n with n suppliers is *incentive compatible* if it is so for the suppliers, thus satisfying $q \leq q(n)$, *and* also for the principal as described next.

Lemma 2 (Incentive-compatible relational contracts) *Any relational contract C^n that contemplates restricted competition and strictly positive quality and that solves the principal's program with commitment $(\mathcal{P}_c)$ is incentive compatible. With open competition, a relational contract C^N with quality $q > 0$ is incentive compatible if and only if $q \leq q(N)$* and

$$s(\hat{q}^*) - \theta\left(\hat{n}^*\right) \geq s\left(q\right) - \theta\left(N\right) - \frac{1}{N}vq, \tag{4}$$

where $\hat{q}^*$ and $\hat{n}^*$ solve the principal's commitment program $(\mathcal{P}_c)$ with the additional constraint of restricted competition $\hat{n}^* < N$.

With restricted competition, i.e. $n < N$, the principal can always credibly substitute a deviating supplier at no cost, with one of the $N - n$ suppliers initially excluded from the pool. Furthermore, when n and q are the maximizers q^*, n^* of the principal's commitment program $(\mathcal{P}_c)$, deviating to a different n (e.g. $n = N$) cannot be optimal either, because suppliers would then begin providing nil quality. Hence, when procuring with restricted competition, the principal's incentive compatibility is clearly satisfied and optimal procurement is simply the solution of the principal's program $(\mathcal{P}_c)$, with the additional constraint $n < N$.

When instead procuring with open competition, i.e. $n = N$, constraint (4) guarantees that the principal is credible in excluding a cheating supplier. By retaining a supplier who cheated, she will continue procuring with N suppliers and a per-period payoff

$$s\left(q\right) - \theta\left(N\right) - \frac{1}{N}vq,$$

which accounts for the fact that in any future period, there is a probability $1/N$ of the auction winner being precisely that supplier who will continue to deliver nil quality. By excluding a supplier who cheated, the principal will instead obtain a per-period payoff associated with the solution of problem $(\mathcal{P}_c)$ with restricted competition, i.e. $s(\hat{q}^*) - \theta\left(\hat{n}^*\right)$. In fact, once a deviating supplier is excluded, procurement will take place under restricted competition, from the point of view of all other suppliers. When there exists no q that satisfies (4) and (1), with open competition the principal can only procure nil quality.

The principal's optimal relational contracts thus solves the following optimization program:

$$(\mathcal{P}) \qquad \begin{array}{l} \max\limits_{N \geq n,\ q(n) \geq q \geq 0} V(n, q) \\ \text{st. (4) when } n = N. \end{array}$$

Note that the principal can always guarantee herself the payoff V_0^* of the zero quality equilibrium, which is in fact equal to $V(N, 0)$ since $q = 0$ trivially satisfies all constraints. The principal can obtain some positive quality with open competition as long as she can exclude from future auctions a supplier who fails to deliver the expected quality. She can further increase the level of procured q by restricting competition to a smaller pool of suppliers $n < N$ from the very beginning, i.e. by adopting restricted competition. Indeed, reducing competition increases suppliers' expected rent in two complementary ways: by increasing the informational rent from winning (because $\pi(n)$ is decreasing in n) and by increasing the probability/frequency of winning (because $\beta(n)$ is decreasing in n). The cost of restricted competition is that the price paid and the associated cost $\theta(n)$ faced by the principal are increased with a smaller n.

The principal's program ($\mathcal{P}$) shows a fundamental trade-off between quality and competition. Since the expected informational rent $\beta(n)\pi(n)$ is decreasing in n and, as a consequence, the maximum enforceable quality $q(n)$ too, to procure a higher quality q and at the same time ensure incentive compatibility, the principal may be obliged to severely restrict the number of invited suppliers, with the highest enforceable quality obtained by negotiating with a single supplier (i.e. $n = 1$). In other terms, more intense competition allows the principal to procure at a lower cost, but at the same time it also reduces the non-contractible quality q. The next proposition illustrates how to address this fundamental trade-off.

Proposition 1 (Optimal relational procurement) *There exists a decreasing function $v(N)$ that converges to 0 so that the principal's optimal relational procurement is as follows:*

- *When $v < v(N)$, i.e. quality is not very important for the principal, then $n^* = N$, i.e. open competition is optimal with either a strictly positive quality $q^* = \min\{q_{FB}, q(N)\}$ if (4) is satisfied, or $q^* = 0$ otherwise.*

- *When quality is important for the principal, i.e. $v \geq v(N)$, restricted competition is optimal with a number of invited suppliers $n^* = n(v) < N$, weakly decreasing in v, and with quality $q^* = \min\{q_{FB}, q(n^*)\}$.*

When the value of quality is limited there is no reason to restrict competition, which would only be meaningful for the principal to possibly increase quality. In this case, open competition is optimal. Then either procurement delivers a positive quality if it is incentive compatible (satisfying both and (4) and (1)), or the principal prefers to procure with maximal (open) competition and accept nil quality.

When instead quality matters, the pool of n actual competitors must be restricted, otherwise the procured quality would be too low. How much the principal wants to restrict competition depends on the trade-off between reduced screening efficiency and higher quality. The larger v is the more this trade-off is resolved sacrificing competition to increase procured quality, and this is why, overall, the optimal number of invited suppliers is decreasing in v.[17]

Note that if the number of potential suppliers is very limited, i.e. N is very small, then open competition may still be optimal even if the principal values quality (indeed the function $v(N)$ is decreasing in N). This is because the procurement cost increase caused by excluding even just one supplier i.e. $\theta(N-1)-\theta(N)$, can be very large when N is small.

The proposition is reminiscent of the efficiency wage models and related applications to procurement (Shapiro and Stiglitz 1984, Kim 1998, Board 2011). However, it differs substantially as it also suggests that restricting competition – with the cost-efficiency loss it entails – is not always necessary to obtain (positive) non-contractible quality. This is indeed the case when the value of quality v for the principal is intermediate and the number of potential suppliers N is also not too large. As long as some discretion is present, involving the possibility to switch to restricted competition, positive levels of non-contractible quality can be elicited with open competition, and when quality is not too important this may suffice.

Summarizing, due to incentive compatibility constraints, restricted competition is more likely to be optimal for the principal, the more valuable non-contractible quality is (the higher v), the larger the number of potential suppliers (the larger N), the smaller the cost heterogeneity across suppliers (hence, the informational rent $\pi(n)$), and the smaller the discount factor δ.

Using open competition for standardized and easy-to-contract transactions (where v is nil or small), and restricted procedures or negotiations for core supplies and complex services, is common practice in private procurement. Public procurement laws also typically admit restricted procedures and negotiations when the object of the transaction is complex or difficult to contract upon. In the US and Europe, the possibility of using simplified procedures for public procurement, like restricted auctions, that allow to take past performance into account and to build relationships

[17] The principal may guarantee future rents, and thus high quality, using a minimum price. However, a minimum price is suboptimal to guarantee future rents because inefficient firms may at times win, obtaining small rents and reducing procurement efficiency.

has been long debated. Although the discretion can be abused for relational contracts whose objects are bribes rather than quality, recent evidence shows that the ability to use restricted auctions with invited bidders is often associated with better performance.[18]

That a large number of suppliers in the market (N) and a decrease in their cost heterogeneity should, ceteris paribus, induce more use of restricted competition has strong cross-industry implications. The conclusion appears also to be consistent with what has been observed in the last decades with IT procurement practices. Globalization has involved a drastic fall in trade barriers and has increased the number (N) of suppliers around the world that each buyer can access. At the same time, the increase of competition also associated with globalization has driven out of business the most inefficient suppliers, containing or even reducing market heterogeneity in terms of cost. In the light of our previous results, it is easier to understand why many buyers have reacted to this process by restricting attention to smaller networks of regular suppliers (see, for example, Bakos and Brynjolfsson, 1993).

The last comparative statics on δ can be interpreted in terms of contract duration and implied frequency of interaction.[19] A more intense use of restricted competition should be observed in environments where the average (explicit) contract duration is long, implying a low frequency of interaction between buyer and suppliers (when the unitary time interval is long, frequency of interaction is low making the discount factor δ small).

5 Optimal procurement with collusion "risk"

Proposition 1 shows that where quality is important for the principal, limited enforcement requires reducing the number n of invited suppliers, and this requirement is stronger the higher is v. However, it is well known that repeated interaction between a restricted number of long-lived suppliers tends to foster collusion among them (as well as increasing the frequency of interaction which here can be interpreted as an "endogenous" reduction of δ). This means that a principal relying on restricted competition to control non-contractible quality may risk inducing suppliers to collude, which is more than just a possibility, as illustrated in the Introduction.

[18] For Europe, see Coviello et al. (2017) and Chever et al. (2013), for the US see Gil and Marion (2011) and Kang and Miller (2016), who point to positive effects of relationships and restricting competition.

[19] In Calzolari and Spagnolo (2009), we explicitely considered contract duration and showed that there are complementarities between shortening existing contract duration and reducing the number of invited agents.

We consider the possibility of suppliers' collusion in the simplest way, focusing on the case in which open competition is dominated according to Proposition 1. We take the view that collusion takes place whenever it is viable, i.e. it is incentive compatible. Collusion takes the form of a "stochastic" bid rotation driven by the efficiency cost parameters θ: the most efficient supplier is awarded the contract at the reservation price r and all the others in the pool of n invited suppliers either refrain from bidding or submit losing bids.[20]

Enforcement and collusion: a trade-off. Let $\tilde{\pi}(n)$ denote supplier i's expected rent when a cartel among the n suppliers is in place, he is the most efficient cartel member, and he chooses not to deviate, i.e. $\tilde{\pi}(n) = r - \theta'(n)$. For collusion to be sustainable, the second-most efficient supplier, i.e. the one with the strongest temptation to deviate, must not prefer to undercut the most efficient one. If this supplier does not deviate, he can expect the future collusive profits. Otherwise, by deviating he gets an immediate gain that we indicate with $D(n) \geq 0$, but then collusion breaks down and all suppliers will compete from then on (grim trigger strategy). This deviation is dominated if the following holds:

$$\beta(n)\tilde{\pi}(n)\frac{\delta}{1-\delta} \geq D(n) + \beta(n)\pi(n)\frac{\delta}{1-\delta}.$$

Clearly, a supplier who deviates from the collusive agreement may also consider the possibility of cheating with a nil quality. Such a deviation is dominated if

$$\beta(n)\tilde{\pi}(n)\frac{\delta}{1-\delta} \geq \psi(q). \tag{5}$$

Hence, collusion is viable if the following incentive compatibility constraint is verified:

$$\beta(n)\tilde{\pi}(n)\frac{\delta}{1-\delta} \geq D(n) + \max\{\beta(n)\pi(n)\frac{\delta}{1-\delta}, \psi(q)\}. \tag{6}$$

What is the effect of a smaller number n of suppliers in the pool on cartel stability? First note that, for a given collusive bid, a smaller n reduces $D(n)$, since the second-most efficient supplier becomes, on average, less efficient and his gains from deviation are thus smaller. Second, although both the l.h.s. and the second term in the r.h.s. of (6) decrease with n, the first falls more than

[20] We do not consider partial collusion involving fewer than n agents. To simplify exposition, we also assume the cartel can implement high winning price for any realization of costs. Alternatively, for a sufficiently low realization of costs, the cartel may contemplate temporary reversion to competitive bids. This possibility would not qualitatively alter our results.

the second. This is immediate if the deviating supplier also cheats on quality, that is, when the max in r.h.s. of (6) is $\psi(q)$ which does not depend on n. When this is not the case, the cartel stability condition (6) becomes

$$\beta(n)[\tilde{\pi}(n) - \pi(n)]\frac{\delta}{1-\delta} \geq D(n),$$

where the left-hand side increases in n. In fact, a smaller size n of the pool of invited suppliers increases both the probability of each of them of being the most efficient $\beta(n)$ and also the difference $\tilde{\pi}(n) - \pi(n)$, which is in fact equal to $r - \theta(n)$. This reasoning immediately leads to the following.

Proposition 2 (Enforcement vs. collusion) *There is a trade-off between enforcement and collusion: by reducing n in order to increase non-contractible quality q, the principal may facilitate collusion among suppliers.*

This is a general trade-off that is clearly relevant in many other frameworks that share the ingredients of competition among suppliers and the need to give them the long-run incentives to perform non-contractible tasks. This trade-off is general also because it applies to other factors that tend to facilitate cooperation besides the number of suppliers. For example, a shorter contractual duration, here captured with a higher δ, facilitates the enforcement of quality by increasing the frequency of interaction and the speed at which a non-performing supplier is punished, but it also facilitates collusion between suppliers, by increasing the speed at which a cartel defection is punished by other cartel members.

Optimal relational contracts with collusion. How does the principal address this trade-off between enforcing non-contractible quality and suppliers collusion? The following Lemma provides a first important result.

Lemma 3 (Quality with collusion) *Any level of non-contractible quality q that is enforceable with competing suppliers, is enforceable with colluding suppliers.*

This is a simple consequence of the fact that the stability of collusion (i.e. constraint (6)) implies that the expected present value of profits with collusion is greater than with competition. Quality q is enforceable with collusion if $\beta(n)\tilde{\pi}(n)\frac{\delta}{1-\delta} \geq \psi(q)$. This condition is implied by q being enforceable with competition, i.e. $\beta(n)\pi(n)\frac{\delta}{1-\delta} \geq \psi(q)$, because a necessary condition for collusion to be viable is $\tilde{\pi}(n) \geq \pi(n)$. Hence, any supplier, when deciding whether or not to deliver q, knows that the cost of cheating on quality is greater than under competition, since the profits at stake are larger. We can now state the following.

Proposition 3 (Optimality of collusion) *When quality is sufficiently important for the principal (high v), optimal procurement involves few colluding suppliers (small n) or negotiation ($n = 1$), and a higher quality q^* than without collusion.*

When quality is important to the principal, limiting competition may become desirable above and beyond the principles of Proposition 1. Not only may the principal optimally want to restrict the number of invited suppliers, she can improve quality further by inducing them to cooperate in cartels or consortia, because with collusion the suppliers' expected profits increase, which ultimately makes better quality attainable. By freezing competition, collusion relaxes the trade-off between the number of suppliers and rents needed to enforce non-contractible quality, thereby allowing the principal to increase quality.

Clearly, if the principal can restrict competition to the case of single-party negotiation, and this is optimal (i.e. $n^* = n(v) = 1$; see Proposition 1 and the program for negotiation (19) in the Appendix), the distinction between competition and collusion vanishes. However, there are several factors that make procurement with colluding suppliers the best way to procure when quality matters. There are cases in which a principal cannot restrict attention to a single supplier, for example in public procurement.[21] The principal may have a general concern for efficiency in production, in which case negotiating with a single supplier delivers the lowest possible expected efficiency. This may be the case for example with a public buyer, and the concern for efficiency will naturally emerge when the principal can procure using ex ante and ex post transfers as we will show in Section 6.

Bilateral negotiation is also dominated when the supplier may be subject to unexpected inability to procure, a fact that exposes the principal to the risk of a costly unplanned halt in procurement, whenever she negotiates with a single supplier. To reduce or eliminate this risk, the principal can ensure supply by relying on second-sourcing. Assume that the adverse event (is observable and) takes place with probability α, in which case the unique supplier would be unable to procure. Facing this risk of no procurement—the costs of which is at least v_0—the

[21]In most countries' small-scale public procurement and that of international organizations like the United Nations, accountability rules require obtaining competitive offers from some minimum number of potential suppliers.

principal may allocate two contracts. The first-source contract is as in the previous analysis and contemplates the possibility that, with probability α, the supplier cannot procure. The second-source contract is only executed with probability $1-\alpha$, when the first supplier cannot procure. Since the principal will never optimally allocate the two contracts to the same supplier, to avoid exposing herself again to the risk of no procurement, dual-sourcing corresponds to a multi-unit auction where suppliers are not allowed to win both contracts.[22] The cost of second-sourcing for the principal is that the winning bids tend to be higher because of the multi-units. On the other hand, second-sourcing guarantees procurement even in the case that the adverse event realizes. When the cost of halting procurement is large enough, and quality matters, procurement with colluding suppliers dominates negotiation.

As stated in the Introduction, what we refer to here as "collusion" can be also taken to stand for any cooperative and self-sustaining agreement among suppliers, such as consortia, and other forms of joint bidding like joint ventures. In this case, the principal *de facto* deals with a single consortium which is stable because of the incentive compatibility constraint (6), regardless of any legal obligation among agreeing supplier. Importantly, the internal stability of the consortium could not be entirely replicated with a formal contract precisely because of the non-contractible dimensions. Moreover, in cases in which consortia are barred by law, inducing suppliers to collude implicitly is indeed a way to get around the prohibition and recover the benefits of supplier cooperation and higher quality.

One may wonder that collusion induces entry of competitors ready to underbid the high collusive prices thus disrupting the existing and collusive relational contracts. However, note that the level of competition is directly controlled by the buyer who may impede entry. Moreover, in return for the large collusive rents the buyer asks for a very high quality, which even a very efficient entrant may be unable to offer at a reduced price.

It is important to stress that the optimality of suppliers collusion for the buyer does not imply optimality for society. As mentioned in the Introduction, in Japan for many years public buyers have been centrally organizing cartels among potential suppliers, a practice called *Kansei-Dango* (Hayashi, 2016). This practice has been tolerated by Japanese public procurement regulators and antitrust authorities, even though it was known to often being accompanied by bribes and related corrupt practices, because it was explicitly considered the way to enforce quality of public

[22] With at least three competing suppliers, the buyer's selection mechanism can be a uniform-price auction which is efficient in this environment and involves truthful bidding.

procurement, much like our Proposition 3 suggests. Suppliers collusion may have negative effects not explicitly accounted for in the model, such as rent-seeking and associated bribes, or negative externalities on other buyers, a possibility we will elaborate on in Section 8.

6 Procurement with ex ante and discretionary transfers

So far we have assumed that the only admissible transfer between the principal and the suppliers is that governed by the explicit contract attributed with the auctions, i.e. the payment to the winning bidder. This is a framework typical of public procurement and large bureaucratic organizations, where accountability reasons tend to restrict the type of admissible transfers, particularly discretional ones like bonuses and bonds conditional on non-contractible performance. In private procurement and employment relationships, instead, these restrictions are often less relevant. Here we investigate if and how enriching the set of available however affects our results.

With respect to a date t, transfers can be paid *ex ante* before suppliers learn their cost and bid, *interim* such as the payment to the auction winner, and *ex post* after delivery of procured goods or services. An *ex ante transfer* $w \in \Re$ could be seen as a fee $w > 0$ for participating in the selection process ("participation fee") or a fixed wage $w < 0$ to compensate the n suppliers in the pool for being available to the principal ("availability wage"). At the end of a contract, the principal and the supplier may also exchange an *ex post transfer* $B \in \Re$ which can be a bonus $B > 0$ that the principal may discretionally decide to pay, or a performance bond $B < 0$ posted by the supplier at the beginning of the contract that the principal may discretionally decide not to return.[23] The possibility to rely on competing suppliers brings about some interesting features of ex ante and ex post transfers in relational contracting that significantly stand out with respect to the case of a single available supplier, i.e. $N = 1$, typically analyzed in the literature.

Several papers have illustrated the desirability of discretionary ex post payments as a disciplining device when a single supplier is available (e.g. MacLeod and Malcomson, 1989 and Baker et al., 1994). Interestingly, with competition any promised bonus is "competed out" at the auction stage even if it is paid ex post. In fact, any ex post transfer B that the principal announces to pay to the winner after the auction induces suppliers to reduce their bid by an equivalent amount, so that the winning bid becomes

$$b' = \theta(n) + \psi(q) - B.$$

[23] Bonuses and bonds exchanged between the principal and an agent that is not the current supplier are irrelevant.

In the end, the principal does not pay B and the bonus becomes a *de facto* bond in the hands of the buyer, which the latter can decide to withhold discretionally. With competing suppliers bonuses/bonds become an effective and inexpensive out-of-equilibrium threat to the current supplier.

This is not the case, of course, for ex ante transfers w. Although these transfers cannot be used as threat, they allow the principal to discipline suppliers by directly increasing their expected rents that now become $w + \beta(n)\pi(n)$. Suppliers' incentive compatibility constraint with these ex ante and ex post transfers then becomes

$$[w + \beta(n)\pi(n)] \frac{\delta}{1-\delta} \geq \psi(q) - B. \tag{7}$$

With competing suppliers, these two types of transfers have drawbacks as well. When positive, the ex ante transfer w is expensive because it is paid to all the n suppliers in the pool of competitors, with a cost of $n \times w$. This implies that the principal will set these transfers at their minimum and the suppliers' incentive compatibility constraint (7) is optimally set as an equality.[24] Ex post discretionary payments, instead, cannot be used with restricted competition. Since procured quality is observable in the principal-supplier dyad only, when $n < N$ the principal can always renege on a bonus claiming that the supplier under-performed and replace him at no cost with another supplier among those $N-n$ outside the pool.[25] When the principal chooses to procure with open competition instead, a discretionary payment conditional on quality may become credible. By reneging on the bonus, the principal will face the cost of contracting with fewer $N-1$ suppliers in the future. Formally, for a bonus of size B to be credible, it must be lower in size than the reduction of procurement value to the principal if he decides to renege on the bonus, i.e.

$$[V(n,q) - \underline{V}]\, \delta \geq B, \tag{8}$$

where

$$V(n,q) = [s(q) - \theta(n) - wn] \frac{1}{1-\delta}$$

and $\underline{V}$ is his continuation payoff when the current supplier can no longer be employed because the principal has reneged on the bonus. Clearly, when $n < N$ we have $\underline{V} \geq V(n,q)$ and no strictly

[24] This is also the case when $w < 0$ (a participation fee), since the principal wants w to be as large as possible in absolute value so that (7) binds at the optimum.

[25] This observation is related to efficiency-wage equilibria in the labor market where workers are on the "long side" and firms on the "short side" of the market. As shown in MacLeod and Malcomson (1998), bonuses are not credible in this case either.

positive bonus is credible. When instead $n = N$, then the l.h.s. of (8) is positive and represents the upper-bound for a credible bonus. We can then state the following.

Proposition 4 (Procuring with ex ante and ex post transfers) *When ex ante and ex post transfers are available, relational contracting is such that when the value of quality v is high, the principal optimally procures with restricted competition, a number of suppliers n^* decreasing in v and q^* that maximize*

$$[s(q) - \theta'(n)]\frac{1}{1-\delta} - \frac{n}{\delta}\psi(q), \tag{9}$$

$B^* = 0$ and w^* negative ("participation fee") for relatively low v or positive ("availability wage") for large v.

When instead the value of quality v is low, open competition is optimal with $B^* = \psi(q^*)$, $q^* = q^{FB}$ if N is small, otherwise q^* is bounded from above by

$$\psi^{-1}\left[[\theta'(N-1) - \theta'(N)]\frac{\delta}{1-\delta}\right] \geq q^*, \tag{10}$$

and $w^* < 0$ ("participation fee").

As with interim transfers only (Section 4), open competition turns out to be optimal and fully efficient when the number of potential suppliers N is relatively small. It is also interesting to note that in this case, any supplier in the pool of competitors must pay a participation fee $w^* < 0$. This is the case because suppliers' incentives (7) are entirely governed by the bonus: since a bonus is never actually paid with competing suppliers, for any q the principal optimally sets it at $B = \psi(q)$ and directly controls quality. Ex-ante transfer w can then be used to extract suppliers' expected informational rents.

However, if N is large, relying on open competition requires a reduction in quality which is now dictated by the principal's incentive compatibility constraint (8). The reduction of the principal's payoff when wrongly retaining the bonus, i.e. the left-hand side of (8), is ultimately proportional to the increase in the expected procurement costs, $\theta'(N-1) - \theta'(N)$, when procuring with $N-1$ instead of N suppliers as in constraint (10) (which is a re-writing of (8)). For a large N the difference $\theta'(N-1) - \theta'(N)$ is small, which limits the bonus and thus the quality, and restricting competition becomes then preferable. Clearly, the same reasoning holds when the value of quality v is large.

With restricted competition, no bonus is credible and n and q are determined maximizing (9) where the term $n\psi(q)/\delta$ is the rent that the buyer pays to all invited suppliers with the transfer

w. Since this cost is increasing in both q and n, it implies that when v is high so that the principal prefers a high quality, she will procure with fewer suppliers, reducing n. Finally, we note that when quality is very valuable and then large, the right-hand side in (7) is large as well and the principal sets a positive w^*, as in the case of an employment contract between the principal and the n invited suppliers.

We note that the inability to use the bonus with restricted competition is consistent with the observation that discretionary monetary bonuses or bonds are hardly ever observed in auto part relational procurement, where competition is typically restricted to two or three trusted suppliers, as documented in Calzolari et al. (2016) for the German car industry.

6.1 Collusion "risk"

Can collusion be optimal, as illustrated in Section 5, even if the principal uses ex ante and ex post transfers?

With an optimal bonus $B = \psi(q)$ the principal fully controls suppliers' incentives. Since the benefit of colluding suppliers is the higher rents they can obtain and the associated larger implementable quality, collusion would only induce higher procurement costs with no effect on implementable quality. Procuring with colluding suppliers is thus not a good idea when the principal uses bonuses. However, we already know that when the value of quality is high for the principal she prefers to restrict competition, and we also know that collusion may be desirable because of the higher quality it allows. Hence, it is natural to investigate collusions with restricted competition, in which case the bonus must be nil and different effects kick-in.

Proposition 5 *With ex ante transfers, procurement with colluding suppliers can be optimal and strictly better than negotiation when the principal wants to implement high non-contractible quality.*

Collusion grants higher rents to the suppliers, which allows the principal to implement higher quality independently of the type of available transfers. Hence, for the same reason discussed with interim transfers only, if the principal very much cares for quality, collusion is preferable to competition.

What is new with additional transfers is that collusion can now directly dominate bilateral negotiation even if there is no risk of halting procurement. The reason is that with ex ante transfers, the principal now appropriates the suppliers' surplus and internalizes the actual cost of production. Since negotiation is associated with the highest expected cost, collusion may be preferable because it also grants high quality with relatively low cost of production.

7 Extensions and robustness

In this section we extend the model to multilateral contracting and discuss some robustness checks.

7.1 Multilateral contracting

As in MacLeod and Malcomson (1998), when a deviation occurs, either by the principal (reneging on B) or by a supplier (not providing the agreed upon level of q), the other players are unable the observe identity of the deviator. This is consistent with q being non-contractible, in the usual sense of being only observable by the two parties involved in the specific relationship, and not by third parties, such as other suppliers or a court, as otherwise these third parties could help enforce an explicit contract on q.

However, when discretional monetary payments are allowed, as is typical in private sector procurement, other possibilities are of interest. An alternative assumption we consider here is that, although courts are not able to observe q, other suppliers who participated in the auction can. This means extending our analysis to multilateral and interdependent relational contracts, as analyzed in Levin (2002).

In an employment environment, suppliers work simultaneously and possibly in teams, as in Levin (2002), thus naturally observing each other's behavior. When suppliers compete however, such as with procurement or consulting, they rarely work simultaneously. Instead they supply in turn and sequentially and have less occasions to observe competitors' behavior, unless the principal plays an active role. To explore this possibility imagine that, at some cost and irreversibly, the principal can form and maintain information sharing among suppliers so that the quality effectively procured by each supplier becomes common knowledge among all the n invited suppliers.[26] This assumption squares well, for example, with Toyota's choice to invest considerable

[26] For simplicity we assume that the cost of sharing information is a one-off set-up cost, and consider how our game is affected after this cost is undertaken (being sunk, the cost itself will not affect the continuation game). The possibility to endogenously affect observability in relational contracts has been explored by Kvaloy and Olsen (2009).

resources in creating and maintaining an organized supplier network, the BAMA, as discussed in the Introduction, to facilitate information sharing with and among suppliers.

Consider first the simpler case of *open competition* with $n = N$. Information sharing and multilateral contracting only affect suppliers' out-of-equilibrium behavior: a multilateral relational contract C^N_{mult} now prescribes that all the N suppliers start punishing the principal as soon as she deviates against any of them. The coordinated punishment substantially reduces the continuation payoff of the principal after her deviation. As a consequence, the upper-limit for an implementable bonus B increases considerably and a higher level of quality can be sustained. Note that the principal's benefit from introducing multilateral contracting and obtaining higher quality is in principle larger in our sequential competitive environment than in teams, as in Levin (2002). The reason is that when suppliers act simultaneously, all bonuses should be paid at the same time and the principal can renege on *all* suppliers' bonuses at once.[27] In our set up, instead, while the multilateral contract strengthens the punishment phase for a deviating principal, it does not affect the principal's maximal gains from an optimal deviation, i.e. retaining the bonus of just one supplier.

When the principal relies on *restricted competition*, the effect of multilateral relational contracting is less immediate. The principal may find it impossible to costlessly replace a supplier after reneging on the bonus as other suppliers observe her deviation, and this may make bonuses credible also with restricted competition.

Proposition 6 (Multilateral contracting) *Multilateral relational contracting is of no value to a principal who procures with restricted competition. With respect to bilateral contracting, it strengthens the scope for open competition, the use of bonuses, and it increases the level of implementable quality.*

Depending on the cost the principal faces to guarantee information sharing among suppliers, procuring under multilateral contracting may be desirable for the principal as it creates a commitment device and enhances her credibility for discretionary bonuses. When this is the case

[27] This is similar to collusion with multimarket contacts. The net gain from multilateral contracting emerges from smoothing asymmetries, i.e. from pooling the incentive constraints of the different bilateral relations into a single joint constraint that optimally reallocates scarce enforcement power across them. In a simultaneous environment, gains from a multilateral contract may also come from payoff interdependence, for example generated by technological externalities (Levin 2002, Spagnolo 1999), which are less relevant in our sequential supply environment.

the principal tends to manage the non-contractible quality with the bonus and prefers to increase the level of competition, thus making collusion less desirable for the principal. However, an information-sharing device that helps suppliers cooperate in punishing a principal's deviation will typically also help suppliers collude and cooperate against the principal. Interestingly, these observations point to the fact that when multilateral contracting is optimal, collusion may emerge even if it is now disliked by the principal, another type of the enforcement-collusion trade, as in Proposition 2.

7.2 Robustness

Although we assumed fully informed suppliers, the drivers of our results also hold qualitatively in a more complex environment with privately informed suppliers. The analysis of competing suppliers can be extended to asymmetric information almost without modification because, applying the revenue equivalence theorem, all standard auction formats would remain equivalent for the principal. With colluding suppliers, significant complexities emerge, as discussed in the literature on collusion with repeated auctions (see Skrzypacz and Hopenhayn, 2004). As the cartel members are privately informed, the efficiency properties of the cartel would be weakened. However, what matters for our results on collusion is the fact that the latter is incentive-compatible. Since this necessarily implies that (equilibrium) expected profits with collusion are greater than with competition, ultimately, enforceable performance is higher than with competing suppliers. Furthermore, the comparison between procuring with colluding suppliers and negotiations is also qualitatively unaffected. On average, the supplier selected out of the many (colluding) suppliers is more efficient than the single supplier with negotiation.

In our analysis we followed Levin (2003) and other previous work on relational contracts with asymmetric information in assuming that suppliers' efficiency is IID. This is a simplification, and in some cases cost persistency, as in Malcomson (2016) for example, may be more realistic. In this case, the principal would learn over time from auctions, and the cost of dismissing a cheating but efficient supplier would be higher than in our environment. Setting aside the complications of such a model, exclusion could be less of a deterrent, suggesting an intrinsic trade-off between efficiency and performance. On one hand, less efficient suppliers would be aware that they can be readily discarded and replaced, and this provides the right incentives. On the other hand, more efficient suppliers know that the principal would be reluctant to discard them and so will be less disciplined in providing high non-contractible performance.

Recent theoretical papers on relational contracts with subjective performance measures (Levin 2003, MacLeod 2003 and Fuchs 2007) emphasized that the realized performance observed by the principal may be subject to noise, so that the principal and the supplier have private information on what they observe. A common theme in this environment with imperfect private monitoring is that in order to induce the principal to report perceived performance truthfully, the optimal contract must make the principal indifferent between reporting different performance levels. In relationships with a single supplier, this tends to induce inefficiencies and the optimal relational contract requires "money burning," which is used by the principal to penalize a non-performing supplier without gaining anything itself. Interestingly, with multiple suppliers, as in our environment, when restricted competition emerges the principal can punish a non performing supplier by replacing him and, what is more, without directly gaining from this move (because suppliers are all ex ante equal and the gain goes instead to another supplier). This maintains incentives for truthful reporting on the side of the principal, and we expect that analogous results in the cases of restricted competition may obtain.

8 Conclusion

We have shown how a principal optimally procures non-contractible quality in a recurrent environment when several suppliers are available. To incentivize the provision of non-contractible quality, the principal must rely on suppliers' future rents which are decreasing in the level of competition. We show that optimal procurement requires restricting competition to a smaller set of competing suppliers when the principal has a strong preference for quality, and open competition among all available suppliers otherwise. This pattern applies independently of the principal's ability to employ ex ante and ex post discretionary transfers. However, restricting competition (as well as reducing the length of the contracts) exposes the principal to the risk of collusion, pointing to a general trade-off between relational enforcement and collusion. We have shown that, unexpectedly, collusion may well be desirable for the principal under some conditions, and in particular when quality and efficiency matter and/or the principal wants to avoid the risk of procurement being interrupted. When open competition is optimal instead, information sharing may improve procurement by allowing multilateral contracting. However, this practice may facilitate collusion in a case in which, however, it is not desirable for the buyer.

These results seem to square well and explain the huge and rather puzzling auto parts cartels that have been discovered recently operating all over the world, as discussed in Section 2. In the light of our results we find no more puzzling that these cartels lasted so long, despite buyers being large and sophisticated; that many colluding suppliers were controlled by Toyota whose performance did not seem to suffer; and only one buyer Ford sued some of the convicted suppliers.

Our results and the above discussions also help clarify part of the disagreement between the US and Japanese authorities during the heated debate on the US-Japan trade imbalance in the 1990s (Cooper, 2014). The large and growing US trade deficit toward Japan in many industries, and in particular in the automobiles and auto parts industries, translated into forceful political pressure from the US that went as far as threatening defence agreements, and led Japan to remove most of the state-sponsored formal and informal trade barriers that appear to have protected Japanese markets. Our model clarifies how long-term relational contracts with few and regular trusted suppliers had exclusionary effects on potential entrants, even after Japan's trade barriers were lifted. Collusion among suppliers, likely to be present already at that time according to our results, could only reinforce this exclusionary effect. In some documents circulating at the time (e.g. Morita, 1993), the fact that supply relations in *Keiretsus* like Toyota were dominated by cartels among suppliers was almost taken for granted. In the light of that debate, we conjecture that the auto parts cartel may have originated well before the dates currently estimated by competition authorities around the world. The pressure from the US was then also directed at increasing the independence and activity of the Japanese Competition Authority, which was seen as strongly subject to the influence of the MITI and of the government, in turn influenced by the big Keiretsus. The Authority had the power to exempt horizontal agreements from antitrust prosecution and used that power generously with Keiretsus like Toyota, under the political pressure of the government. The surprise of several observers when the auto parts cartel investigation finally started in 2010 is therefore puzzling and confirms the need for more research on the interaction between managerial practices, industrial organization and international trade, in the spirit of the pioneering work of Antras (2003) and Antras and Helpman (2004).

Finally, although favoring the organizing buyer, these collusive practices are likely to have had a negative effect on other car manufacturers buying from the same suppliers, and in particular on US manufacturers that were less versed in lean procurement strategies. An interesting extension of our analysis could account that when suppliers are not exclusive and there are multiple buyers with heterogeneous procurement practices, a buyer using lean procurement may enjoy a double competitive advantage from supplier collusion. On one hand the buyer benefits from higher procured quality, and on the other hand her rivals using arm's-length procurement based on supplier competition face higher prices but no higher quality. While suppliers collusion fixes high prices for all buyers, procured quality is likely lower for buyers who rely more on arm's-length competition. The high relative performance of Toyota we observe in Figure 1 could therefore be due also to this externality undermining competitors, with ambiguous overall effects on consumers and welfare.

References

[1] Abreu, D., (1988), "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica*, 56, 383-396.

[2] Andrews I. and D. Barron, (2016), "The Allocation of Future Business: Dynamic Relational Contracts with Multiple suppliers," *American Economic Review*, Vol. 106, no. 9: 2742-59.

[3] Antras, P. (2003) "Firms, contracts, and trade structure." *Quarterly Journal of Economics* 118, 4.

[4] Antras, P., and Helpman, E., (2004), "Global sourcing," *Journal of Political Economy*, 112(3), 552-580.

[5] Aoki, M., and R. Dore, (1994) *The Japanese Firm: The Sources of Competitive Strength*, Oxford: Oxford University Press.

[6] Asanuma, B. (1989), "Manufacturer-Supplier Relationships in Japan and the Concept of Relation Specific Skill," *Journal of the Japanese and International Economies*, 3(1), pp.1-30.

[7] Baker, G., Gibbons, R. and Murphy, K.J., (1994), "Subjective Performance Measures in Optimal Incentive Contracts", *Quarterly Journal of Economics*, 109, 1125-56.

[8] Bakos, Yannis. and Brynjolfsson, Erik (1993) "From Vendors to Partners: Information Technology and Incomplete Contracts in Buyer-Supplier Relationships," *Journal of Organizational Computing*, 3(3), 301-328.

[9] Bar-Isaac, H., (2005), "Imperfect Competition and Reputational Commitment," *Economics Letters*, 89(2), 167-173.

[10] Board, S. (2011), "Relational contracts and the value of loyalty," *American Economic Review*, 101, 7 (2011), 3349–67.

[11] Bulow, J. and P. Klemperer, (2009) "Why do Sellers (Usually) Prefer Auctions?" *American Economic Review*, 99(4), 1544-75.

[12] Calzolari G. and G. Spagnolo (2009), "Relational Contracts and Competitive Screening," CEPR discussion paper n. DP7434.

[13] Calzolari, G., L. Felli, J. Koenen, G. Spagnolo and K. Stahl (2016), "Trust, Competition and Innovation: Theory and Evidence from German Car Manufacturers," mimeo University of Mannheim.

[14] Chever, L., Saussier, S., and Yvrande-Billon, A. (2013). "The Law of Small Numbers: Investigating the Benefits of Restricted Auctions for Public Procurement," mimeo.

[15] Cooper, W., H. (2014) "U.S.-Japan Economic Relations: Significance, Prospects, and Policy Options," Congressional Research Service report 7-5700, November (www.crs.gov, RL32649)

[16] Coviello, D., A. Guglielmo, and G. Spagnolo, (2017) "The Effect of Discretion on Procurement Performance," forthcoming in *Management Science*.

[17] Doni, N. (2006). "The Importance of Reputation in Awarding Public Contracts," *Annals of Public and Cooperative Economics*, 77(4), 401-429.

[18] Fuchs, W., (2007), "Contracting with Repeated Moral Hazard and Private Evaluations," *American Economic Review*, 97(4), 1432-1448.

[19] Fudenberg, D., D. K. Levine and E. Maskin, (1994), "The Folk Theorem in Repeated Games with Imperfect Public Information," *Econometrica*, 62, 997-1039.

[20] Gil, R. and J. Marion, (2013) "The Role of Repeated Interactions, Self-Enforcing Agreements and Relational [Sub]Contracting: Evidence from California Highway Procurement Auctions," *Journal of Law, Economics, and Organization,* 29:2, 239-77.

[21] Goldfein, S. and J. Keyte (2014) "Auto Parts and Antitrust: A Cautionary Tale," *New York Law Journal,* 252, 90.

[22] Hayashi, S. (2016), "Kansei-Dango and Antimonopoly Law in Japan: The Characteristics of Collusion in the Japan's Public Procurement," in *Competition Policy International*, July.

[23] Helper, S., and R. Henderson, "Management Practices, Relational Contracts, and the Decline of General Motors," *Journal of Economic Perspectives*, 28 (2014), 49-72.

[24] Hörner, J. (2002), "Reputation and Competition," *American Economic Review*, 92(3), 644-663.

[25] Kang, K. and R.A. Miller, (2016) "Winning by Default: Why is There So Little Competition in Government Procurement?" Unpublished manuscript, Tepper School of Business.

[26] Kelman, S. (1990), "Procurement and Public Management." AEI Press, Washington, D.C.

[27] Kim, I. G. (1998) "A Model of Selective Tendering: Does Bidding Competition Deter Opportunism by Contractors?" *The Quarterly Review of Economics and Finance*, 38(4), 907-925.

[28] Kranton, R., (2003), "Competition and the Incentive to Produce High Quality," *Economica*, 70, 385–404.

[29] Kvaloy, O. and Olsen, T. E. (2009) "Endogenous Verifiability and Relational Contracting,"

[30] *American Economic Review,* 99: 2193–208.

[31] Laffont, J.-J. and Tirole, J. (1993), *A Theory of Incentives in Procurement and Regulation*, Cambridge, MA: MIT Press.

[32] Levin, J., (2002), "Multilateral Contracting and the Employment Relationship," *Quarterly Journal of Economics*, 117(3).

[33] Levin, J., (2003), "Relational Incentive Contracts", *American Economic Review*, 93(3), 835-847.

[34] MacLeod, W. B. (2003), "Optimal Contracting with Subjective Evaluation," *American Economic Review*, 92(1), 216-40.

[35] MacLeod, W.B. and Malcomson, J.M., (1989), "Implicit Contracts, Incentive Compatibility, and the Involuntary Unemployment," *Econometrica*, 57(2), 447-480.

[36] MacLeod, W. B. and Malcomson, J.M., (1998), "Motivation and Markets," *American Economic Review*, 88(3), 388-411.

[37] Malcomson, J.M., (1984), "Work Incentives, Hierarchy, and Internal Labor Markets," *Journal of Political Economy*, 92(3), 486-507.

[38] Malcomson, J.M., (2016), "Relational Incentive Contracts With Persistent Private Information," *Econometrica*, 84(1), 317–346.

[39] Milgrom, Paul, and John Roberts, (1993) "Johnson Controls, Inc. - Automotive Systems Group: The Georgetown Kentucky Plant," Stanford Graduate School of Business, Case S-BE-9.

[40] Manelli, A. and Vincent, D.R., (1995), "Optimal Procurement Mechanisms," *Econometrica*, 63(3), 591-620.

[41] Morita, M. (1991) "Structural, Impediments Initiative: is It an Effective Correction of Japan's Antimonopoly Policy?" *University of Pennsylvania Journal of International Business Law*, 12(4), 777-809.

[42] Mukherjee, A. and L. Vasconcelos, (2011), "Optimal Job Design in the Presence of Implicit Contracts," *The Rand Journal of Economics*, 42(1), 44–69.

[43] Rayo, L., (2007), "Relational Incentives and Moral Hazard in Teams," *Review of Economic Studies* 74(3), 2007.

[44] Shapiro, C. & Stiglitz, J. (1984), "Equilibrium Unemployment as a Worker Discipline Device," *American Economic Review*, 74(3), 433–444.

[45] Skrzypacz A. and Hopenhayn H. (2004) "Tacit Collusion in Repeated Auctions," *Journal of Economic Theory,* 114(1), 153-169.

[46] Spulber, D.F. (1990) "Auctions and Contract Enforcement," *Journal of Law, Economics and Organization*, 6(2), 325-44.

[47] Stephan, A., (2016), "The Galvanised Steel Tanks Criminal Cartel Trial," unpublished manuscript.

[48] Stiglitz, J. E., (1989), "Imperfect Information in the Product Market," in R. Schmalensee and R. Willig, eds., *Handbook of industrial organization*, 1. Amsterdam: Elsevier Science Publishers, 771–847.

[49] Taylor, C. and S. Wiggins (1997). "Competition or Compensation: Supplier Incentives Under the American and Japanese Subcontracting Systems," *The American Economic Review*, 87, 598–618.

[50] Troya-Martinez M. and L. Wren-Lewis, (2016), "Delegating Relational Contracts to Corruptible Intermediaries," working paper.

Appendix

Proof of Lemma 1. Sufficiency. Consider first the case with $n < N$ and an auction taking place at t. The relational contract C^n described in the text before the Lemma guarantees sufficiency. In fact, if any supplier prefers to bid and deliver q if it wins, he will prefer to do so at any future auction in $t' > t$. The most efficient supplier in t wins the auction and, by providing q, obtains the expected payoff

$$b' - \theta'(n) - \psi(q) + \beta(n)\pi(n)\frac{\delta}{1-\delta}.$$

If he instead cheats, then he sets $q = 0$ (cheating on other qualities is dominated), is replaced by some other supplier in the group of $N - n$ initially excluded suppliers, and gets $b' - \theta'(n)$ since with none of the other suppliers deviating, he will then be excluded permanently. Hence, if the inequality (1) is satisfied, the supplier has no incentive to cheat on q. Also any supplier i who is not the most efficient in t may bid planning to cheat on quality. Incentives to deviate are strongest in this case for the second-most efficient supplier, who could bid $b' - \varepsilon$ with ε small and positive and set $q = 0$. If he instead bids planning not to cheat, he will not win in t, but can still expect the positive future profits. Hence, he will prefer to bid not planning to cheat if

$$\beta(n)\pi(n)\frac{\delta}{1-\delta} \geq b' - \theta\,(n) \tag{11}$$

which is equivalent to (1) since $b' = \theta\,(n) + \psi(q)$.

Consider now the case with $n = N$ and the following relational contract C^N. If a supplier i deviates by not providing the required q, according to C_i he is excluded from future auctions and the principal and the remaining suppliers switch to restricted competition with $n = N - 1$ and to relational contracts in C^{N-1} that are identical to those in C^N for each supplier except for the excluded supplier i. If any of the remaining $N - 1$ suppliers in the pool deviates in the future, he is replaced by the supplier that was not in the pool in the last period. Anticipating that the other suppliers do not deviate, any supplier knows that being excluded for one period implies permanent exclusion and then (1) is again sufficient to induce the most-efficient, and also all other suppliers, not to deviate.

Finally, note that in both cases for n, suppliers' participation is satisfied since the l.h.s. of (1) is their expected payoff from participating and the r.h.s. is non-negative.

Necessity. Consider any relational contract $\tilde{C}^n$ that contemplates $q > 0$. It must be that (1) is satisfied for $\tilde{C}^n$. Indeed, suppose instead that (1) does not hold. It is immediately clear that the most efficient supplier and also the second-most efficient supplier, have an incentive to deviate. ■

Proof of Lemma 2. The first part of the Lemma is immediate. Since with restricted competition the principal can always exclude a cheating supplier at no cost, the only constraint granting incentive compatibility is that of the suppliers (1).

With $n = N$ this is not the case because by excluding one supplier, the principal may face a reduction of payoff. In particular, when excluding the supplier the principal's payoff is

$$V_{DE} = [vq + v_0 - \psi(q) - \theta(N-1)] \frac{1}{1-\delta}$$

and

$$V_{DD} = \left[\frac{N-1}{N} [vq + v_0 - \psi(q) - \theta(N)] + \frac{1}{N} [v_0 - \psi(q) - \theta(N)] \right] \frac{1}{1-\delta}.$$

when instead the supplier is not excluded, where the term multiplied by $1/N$ accounts for the fact that the supplier who cheated will always set $q = 0$ and win the auction with probability $1/N$.

Note that the cheating supplier may win more often anticipating that he will not face the cost of quality, in contrast to his competitors. However, by doing so the other suppliers would learn that the principal did not punish a cheating supplier and they would then stop procuring the positive quality themselves. This would cancel the advantage of the cheating supplier: he would still win with probability $1/N$ obtaining the same information rent $\pi(N)$, but he would not obtain the extra payoff associated with quality cost-saving $\psi(q)$ since none of the others would provide quality either. Hence, the cheating supplier always prefers to mimic the others, only winning an auction when he is actually the most efficient firm.

When $V_{DE} \geq V_{DD}$, which corresponds to (4), exclusion is credible and C^N is thus incentive compatible. ■

Proof of Proposition 1.

Consider first the case of optimal procurement with restricted competition, i.e. $n \leq N-1$.

If the following is satisfied

$$\frac{\theta\,(N-1) - \theta'(N-1)}{N-1} \frac{\delta}{1-\delta} \geq \psi(q_{FB}),$$

which is equivalent to $q(N-1) \geq q_{FB}$, then the solution of the principal's program is $q^* = q_{FB}$ and $n^* = N-1$. Since the efficient quality q_{FB} is implicitly defined by $\psi'(q_{FB}) = v$, we have here

$$q(N-1) = \psi^{-1}\left(\frac{\theta\,(N-1) - \theta'(N-1)}{N-1} \frac{\delta}{1-\delta}\right) \geq q_{FB} = \psi^{-1\prime}(v)$$

or equivalently

$$v \leq \psi'(q(N-1))$$

and we can identify this first case with

$$v < \underline{v}(N) \equiv \psi'(q(N-1)).$$

Now, assume instead $v > \underline{v}(N)$. Then $q_{FB} \geq q(N-1)$ and at the optimum the constraint $q \leq q(N-1)$ must be binding. The principal's program then becomes

$$\max_{N-1 \geq n} \quad \{s[q(n)] - \theta(n)\} \frac{1}{1-\delta}. \tag{12}$$

If the following holds

$$s[q(N-1)] - \theta(N-1) \geq s[q(N-2)] - \theta(N-2)$$

then it is optimal to set $n^* = N-1$. Since $N \in \Im_+$, linearizing and dividing by a unitary reduction in the number of invited suppliers, i.e. $\Delta n = N-1-(N-2)$, the previous inequality is proxied with

$$[v - \psi'(q(N-1))] \frac{\Delta q(N)}{\Delta n} - \frac{\Delta\theta(N)}{\Delta n} \geq 0 \tag{13}$$

where $\Delta q(N) = q(N-1) - q(N-2) \leq 0$ and $\Delta\theta(N) = \theta(N-1) - \theta(N-2) \leq 0$. Inequality (13) can be equivalently rewritten as

$$v \leq \overline{v}(N)$$

where

$$\overline{v}(N) \equiv \psi'(q(N-1)) + \frac{\Delta\theta(N)}{\Delta q(N)},$$

and

$$\underline{v}(N) \leq \overline{v}(N).$$

In this second case with $\underline{v}(N) < v \leq \overline{v}(N)$, the principal cannot set the first-best quality and the optimal quality is $q^* = q(N-1)$.

When instead $v > \overline{v}(N)$, having $N-1$ competing suppliers is dominated by having only $n = N-2$ in the pool of invited suppliers. In this case, the optimal number n^* of invited suppliers satisfies the following double inequality

$$[v - \psi'(q(n^*))] \frac{\Delta q(n^*)}{\Delta n} - \frac{\Delta\theta(n^*)}{\Delta n} \geq 0 \geq [v - \psi'(q(n^*))] \frac{\widetilde{\Delta q}(n^*)}{\Delta n} - \frac{\widetilde{\Delta\theta}(n^*)}{\Delta n} \tag{14}$$

where $\widetilde{\Delta q}(n) = q(n) - q(n+1) \geq 0$, $\widetilde{\Delta\theta}(n^*) = \theta(n) - \theta(n+1) \geq 0$ and $\widetilde{\Delta n} = n - (n+1) < 0$. The left inequality implies that having n^* suppliers in the pool of competitors is better for the principal than having $n^* - 1$, and the right inequality implies that having n^* is also better than $n^* + 1$. The left inequality in (14) is deceasing in v because $\Delta q(n) \leq 0$ and $\Delta n \geq 0$. The right inequality in (14) is decreasing in v because $\widetilde{\Delta q}(n) \geq 0$ and $\widetilde{\Delta n} < 0$. Hence, a higher v implies a weakly lower optimal number of competitors, i.e. $n^*(v)$ is (weakly) decreasing in v.

Consider now optimal procurement with open competition $n = N$. The optimal quality can be either $q = 0$, or $q = q_{FB}$ or $q = q(N)$. Clearly, for $v = 0$, then $q^* = 0$. If

$$v < \underline{v}(N+1)$$

then q_{FB} is incentive compatible for the supplier and satisfies the principal's incentive constraint (4) if

$$v \geq \frac{N}{q_{FB}} \left[\theta\left(N-1\right) - \theta\left(N\right)\right] \tag{15}$$

which is obtained by (4) using the fact that $\underline{v}(N+1) < \underline{v}(N)$ so that with restricted competition the optimum is at q_{FB} with $N-1$ suppliers. When (15) is violated, then again $q^* = 0$. When $v > \underline{v}(N+1)$ so that q_{FB} is not incentive compatible for the suppliers, then the optimal quality with open competition must be lower, i.e. either $q(N)$ or nil, respectively, when (4) is satisfied ore not.

We can now combine the optimal procurement with open and restricted competition.

When $v \leq \underline{v}(N+1)$, then the optimal quality is either nil or q_{FB} both with open and restricted competition. In both cases, it is optimal to procure with open competition which, for the same quality, reduces costs.

When $\underline{v}(N) \geq v > \underline{v}(N+1)$, there are two cases. If $q(N)$ does not satisfy (4) so that the optimal quality with open competition is $q = 0$, then restricted competition is optimal iff v is sufficiently larger, i.e. $vq_{FB} \geq \theta(N-1) - \theta(N)$. If instead $q(N)$ satisfies (4), then restricted competition is optimal if $v[q_{FB} - q(N)] \geq \theta(N-1) - \theta(N)$, i.e., again if v is sufficiently large since we are considering a case in which $q_{FB} > q(N)$.

When $v > \underline{v}(N)$, then in both open and restricted competition, positive qualities must be $q(.)$. Whenever with open competition $q(N)$ violates (4) so that the optimal quality is nil, then restricted competition is clearly preferable for large v. If instead $q(N)$ satisfies (4), then restricted competition is preferable to open competition when

$$s(q(n^*)) - \theta(n^*) \geq s(q(N)) - \theta(N)$$

i.e.

$$v\left[q(n^*) - q(N)\right] \geq \theta(n^*) - \theta(N)$$

where $q(n^*) > q(N)$ and $\theta(n^*) - \theta(N)$. Again, this is satisfied for a large enough v. ■

Proof of Proposition 2. The proof immediately follows from the argument in the text. ■

Proof of Lemma 3. The proof immediately follows from the argument in the text. ■

Proof of Proposition 3. As stated, we take the view that collusion realizes whenever it is viable. The cartel stability condition (6), written as

$$\beta(n)\tilde{\pi}(n)\frac{\delta}{1-\delta} = D(n) + \max\{\beta(n)\pi(n)\frac{\delta}{1-\delta}, \psi(q)\},$$

implicitly defines a threshold $\tilde{n}(q)$, decreasing in q, such that for any $n \leq \tilde{n}(q)$ collusion takes place.[28] Now consider any $v > \overline{v}(N)$ so that with competing suppliers, optimal procurement contemplates restricted competition, i.e. $n^* = n(v) < N$ where, with a slight abuse of notation, $n(v)$ is implicitly defined by

$$\frac{\Delta q(n)}{\Delta n}\left[v - \psi'(q(n))\right] = \frac{\Delta\theta(n)}{\Delta n}.$$

[28] To avoid extra notation, this expression is written as if the principal kept the same n even if she realized that a cartel would break down. Nothing would change in the arguments below considering a different off-equilibrium relational contract.

For sufficiently large v, we necessarily have $n^* \leq \tilde{n}(q(n^*))$ and the principal's payoff shows a discontinuity at $\tilde{n}(q(n))$:

$$\begin{cases} [s(q(n)) - \theta(n)]\frac{1}{1-\delta} & \text{if } n > \tilde{n}(q(n)) \\ [s(\tilde{q}(n)) - r]\frac{1}{1-\delta} & \text{if } n \leq \tilde{n}(q(n)) \end{cases}$$

where $\tilde{q}(n)$ is the maximal enforceable quality with colluding firms, implicitly defined by $\psi(q) = \beta(n)\tilde{\pi}(n)\frac{\delta}{1-\delta}$. We now compare the values of the principal's payoff on the two branches at the discontinuity point $\tilde{n}$. Collusion is preferable as long as

$$s(\tilde{q}(\tilde{n})) - s(q(\tilde{n})) \geq r - \theta(\tilde{n}) \tag{16}$$

that is when the gain obtained with higher quality (the l.h.s.) is larger than the higher cost of procuring (the r.h.s.). The former is increasing in v (recall $s(q) = vq - \psi(q)$) and the latter does not depend on it, so that for large enough v (16) is indeed verified. ■

Proof of Proposition 4. The principal's payoff with ex ante and ex post transfers is,

$$V(n,q) = [s(q) - nw - E(b') - B]\frac{1}{1-\delta}.$$

For what is stated in the text, any relational contract C^n prescribing $q > 0$ with n invited suppliers is incentive compatible if and only if (7) and (8) are satisfied.

Since $V(n,q)$ is decreasing in w, this transfer is set such that (7) binds, i.e.

$$w = \frac{1-\delta}{\delta}\max\{\psi(q) - B, 0\} - \beta(n)\pi(n).$$

The principal then maximizes

$$V(n,q) = [s(q) - \theta'(n)]\frac{1}{1-\delta} - \frac{n}{\delta}\max\{\psi(q) - B, 0\},$$

subject to (8), i.e.

$$[s(q) - \theta'(n)]\frac{\delta}{1-\delta} - n\max\{\psi(q) - B, 0\} - \delta\underline{V} \geq B.$$

The suppliers' participation constraint is always satisfied because the r.h.s. of (7) is positive.

Now consider open competition. A larger B increases the objective function $V(n,q)$ and, substituting the previous expression for $V(n,q)$ into (8),

$$\left\{[s(q) - \theta'(n)]\frac{1}{1-\delta} - \frac{n}{\delta}\max\{\psi(q) - B, 0\} - \underline{V}\right\}\delta \geq B$$

we note that, as long as $B \leq \psi(q)$, a larger B also helps satisfy (8). Indeed, the marginal effect of a larger B in the l.h.s. of (8) is $n \geq 1$, whilst that on the r.h.s. is simply 1. This shows that the principal optimally sets $B = \psi(q)$ (an even higher B has no effect on V_N and is "costly" in terms of constraint (8)).

Now note that if the principal reneges on the bonus, her outside option $\underline{V}$ is not smaller than the payoff she can get employing $N-1$ suppliers and requesting the same (possibly suboptimal with $N-1$ suppliers) quality q, i.e. $\underline{V} \geq V(N-1,q)$.[29] Since

$$V(N,q) - V(N-1,q) = [\theta'(N-1) - \theta'(N)] \frac{1}{1-\delta}$$

we obtain the condition (10) in the proposition.

Consider now restricted competition. Since by replacing a supplier with any of those $N-n$ in the pool of excluded suppliers, the principal gets $\underline{V}(N-1) = V(n,q)$, (8) implies $B = 0$ and the principal's objective becomes as in (9). Furthermore, the transfer $w = \frac{1-\delta}{\delta}\psi(q) - \beta(n)\pi(n)$ is positive if the optimal quality is high enough. The comparative statics with respect to v with restricted competition directly follows from the objective function (9).

Finally, it is immediately clear that for large enough v, and a fortiori when N is also large, the quality obtained with open competition is severely constrained and the principal prefers restricted competition. ■

Proof of Proposition 5. With ex ante and ex post transfers, the suppliers' incentive compatibility constraint is

$$w + [w + \beta(n)\tilde{\pi}(n)] \frac{\delta}{1-\delta} \geq w + r + \psi(q) - B - \theta'(n) - \psi(q) + B + [w + \beta(n)\pi(n)] \frac{\delta}{1-\delta}$$

which simplifies to

$$[w + \beta(n)\tilde{\pi}(n)] \frac{\delta}{1-\delta} \geq r - \theta'(n) + [w + \beta(n)\pi(n)] \frac{\delta}{1-\delta}$$

and then

$$[w + \beta(n)\tilde{\pi}(n)] \frac{\delta}{1-\delta} \geq r - \theta'(n) + \psi(q) - B.$$

[29] None of the $N-1$ remaining agents is negatively affected by the principal's deviation. Also note that beginning from a candidate optimal contract C^N, the possibility to revert to restricted competition with $n < N-1$ after a principal's deviation is dominated for the principal by contracting instead with all $N-1$ (residual) agents.

Hence, the principal solves the following program,

$$\begin{aligned} &\max_{N\geq n,q,r,w,B} \quad [s(q)-r-B-wn]\frac{1}{1-\delta} \\ [w+\beta(n)(r-\theta'(n))]\frac{\delta}{1-\delta} &\geq \max\{D(n)+[w+\beta(n)(\theta(n)-\theta'(n))]\frac{\delta}{1-\delta},\psi(q)-B\} \\ \left\{[s(q)-r-B-\psi(q)-wn]\frac{1}{1-\delta}-\underline{V}\right\}\delta &\geq B \end{aligned}$$

In particular, focussing on restricted competition, the program becomes

$$\begin{aligned} &\max_{N>n,q,r,w} \quad [s(q)-r-wn]\frac{1}{1-\delta} \\ [w+\beta(n)(r-\theta'(n))]\frac{\delta}{1-\delta} &\geq \max\{D(n)+[w^*+\beta(n^*)(\theta(n^*)-\theta'(n^*))]\frac{\delta}{1-\delta},\psi(q)\} \end{aligned}$$

where n^* is the optimal number of invited suppliers with no collusion and w^* the associated transfer (restricted competition is realized in this case).[30] Since we know from previous analysis that

$$[w^*+\beta(n^*)(\theta(n^*)-\theta'(n^*))]\frac{\delta}{1-\delta}=\psi(q^*),$$

the constraint can be written as

$$[w+\beta(n)(r-\theta'(n))]\frac{\delta}{1-\delta}\geq\max\{D(n)+\psi(q^*),\psi(q)\}. \tag{17}$$

Now we need to consider two cases.

If

$$\psi(q)\geq D(n)+\psi(q^*)$$

then constraint (17) becomes

$$[w+\beta(n)(r-\theta'(n))]\frac{\delta}{1-\delta}\geq\psi(q)$$

which must bind at the optimum for the same arguments discussed in the case of no collusion with ex ante and ex post transfers. The program then becomes

$$\max_{N>n,q} \quad [s(q)-\theta'(n)]\frac{1}{1-\delta}-\frac{n}{\delta}\psi(q) \tag{18}$$

which is as with no collusion except for the fact that the principal now pays the lower cost $\theta'(n)$ of the most-efficient supplier, rather than $\theta(n)$. This immediately shows that when this case is

[30] We realistically assume that the principal notices if a cartel breaks down.

realized, collusion is preferable for the principal. Furthermore, negotiation (i.e. $n = 1$) would lead to the program,

$$\begin{aligned} \max_{q} \quad & [s(q) - r]\frac{1}{1-\delta} \qquad (19) \\ [r - \theta_e]\frac{\delta}{1-\delta} \quad & \geq \quad \psi(q) \end{aligned}$$

where θ_e ($\geq \theta(n)$) is the expected cost of a randomly selected supplier. Since the constraint must bind at the optimum, this program with negotiation becomes

$$\max_{q} \quad [s(q) - \theta_e]\frac{1}{1-\delta} - \frac{1}{\delta}\psi(q). \qquad (20)$$

Comparing the two reduced programs (18) and (20), collusion is optimal when it allows quality and production costs to be better balanced.

If instead

$$\psi(q) < D(n) + \psi(q^*)$$

then the constraint (17) becomes

$$[w + \beta(n)(r - \theta'(n))]\frac{\delta}{1-\delta} \geq D(n) + \psi(q^*)$$

which must bind at the optimum. The program then becomes

$$\max_{N>n,q} \quad [s(q) - \theta'(n)]\frac{1}{1-\delta} - \frac{n}{\delta}[\psi(q^*) + D(n)].$$

The optimal quality of this program is the first-best one, i.e. that maximizing $s(q)$, and the optimal number of invited suppliers n is instead distorted, as usual, for incentive reasons. Hence, comparing this program with competition, it is immediately clear that if the principals cares for quality, inducing collusion is preferable again. This is also the case when comparing with negotiation. ■

Proof of Proposition 6. To restrict the number of possibilities, we assume that as long as one supplier remains in the pool after a deviation by the principal, information concerning the principal's deviation is shared among the suppliers of the possibly new pool and they will all set $q = 0$.

Consider first passive beliefs: the beliefs of the $N - n$ suppliers initially not invited to the pool of competitors about the principal paying any promised bonus are unaffected by decisions to replace some or all of the n suppliers in the pool.

When $N-n \geq n$, a principal who reneges on B can replace at no cost the n suppliers with a new pool of size n of suppliers from the $N-n$ outsiders. This implies that it must be $B=0$.

When $N-n<n$, the principal is able to credibly commit to paying B because the reduction of procuring suppliers from n to $N-n$ is now costly for the principal. However, if the principal finds it optimal to offer a bonus, then restricted competition is suboptimal and in this case $N-n<n$ reduces to $n=N$. To see this, note that the cost for the principal of reneging on B is increasing in n because fewer $N-n$ suppliers would be left available for future procurement, which implies that a larger bonus can be implemented with an even larger n. Moreover, as we have shown with bilateral contracting, when the bonus is credible the principal optimally sets $B=\psi(q)$ (as for open competition in Proposition 4) so that the supplier's incentives are managed and completely guaranteed by the bonus. It is then optimal to increase n up to N also to reduce procurement costs.

Hence, a $B>0$ is only optimally associated with open competition, as with bilateral contracting described in Proposition 4.

To see that this is also the case with non-passive beliefs, assume now that suppliers in the group of $N-n$ outsiders try to interpret a reshuffling of the pool of invited suppliers in terms of deviations of either the principal or the suppliers. When outsiders observe a principal replacing all the n invited suppliers at once, they may believe that the principal is trying to escape a multilateral punishment for a deviation and cannot be trusted in the future.

The principal could conceive of a relational contract C^N_{mult} according to which if a supplier does not provide the expected quality q then all the n suppliers in the pool are replaced. This would have no effect on the supplier's incentive compatibility constraint and would make it impossible for outsiders to know whether a reshuffling of the pool of the n invited suppliers had been induced by a principal's or a supplier's deviation. However, when $N-n>n$ with this contract the principal would have to set $B=0$. She can instead make a $B>0$ credible also in this case by sticking to the type of relational contract contemplated so far, according to which only the current supplier is replace, if needed. With this contract and non-passive beliefs, the principal then prefers to increase n as much as possible, since this allows procurement costs to be reduced with no effect on q, which is governed by B as in the previous analysis. ■