Relational Contracts and Competitive Screening*

Giacomo Calzolari† Giancarlo Spagnolo‡

First draft: November 2004
This version: June 20, 2009

Abstract

We study the tension between competitive screening and contract enforcement when a principal repeatedly trades with one among several agents, moral hazard and adverse selection coexist, and non contractible dimensions are governed by relational contracting. We simultaneously characterize optimal relational contracts and competitive screening policies which are interdependent. When non contractible dimensions are important, the principal optimally restricts competitive screening to a subset of 'loyal' agents giving up performance bonuses, and negotiates with one agent an indefinitely renewable contract when such dimensions are crucial. Explicit contract duration is also optimally reduced to improve enforcement. However, these policies facilitate cooperation among agents inducing an additional trade-off between reputational forces and collusion. When non contractible dimensions are very important this last trade-off may disappear, as collusion allows to enforce higher performance more efficiently. *(131 words)*

*Previous versions of this paper circulated under the title "Reputation and Collusion in Procurement". We have benefited from comments from Pierpaolo Battigalli, Pier-André Chiappori, Mathias Dewatripont, Jonathan Levin, Sergei Izmalkov, Nicola Doni, S. Mitraille, A. Nicolò, Yossi Spiegel, Daniel Traca, Jeanine Thal and Paola Valbonesi. Thanks also to audiences at ESSET 2005 in Gerzensee, ESWC 2005 in London, EARIE 2005 in Porto, CEPR 2005 Workshop on Competition Policy in Brussels, the "Public Services and Management" 2006 Conference in Toulouse, the ESF 2006 Workshop "Designing Partnerships Between Government and The Private Sector" in Bristol, the II IPPC 2006 in Rome, the 2007-Asset meeting, the 2008 CSEF-IGIER Symposium in Capri, ECARES 2009 Conference on Competition in Infrastructure Procurement, and seminar participants at Universities of Berlin (Humbolt), Bocconi, Bologna, Brescia, CERGE-EI Prague, East-Anglia, EUI Florence, Koc, KTH-Stockholm, Mannheim, Padua, Rome 'Tor Vergata', Naples 'Federico II', and the Stockholm School of Economics.

†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.*

JEL Classification Numbers: C73, L14, D86, L41

Keywords: Auctions, Collusion, Cooperation, Contract duration, Implicit contracts, Incomplete contracts, Loyalty, Non-contractible tasks, Procurement, Screening, Relational contracts, Reputation.

1 Introduction

Non-contractible dimensions are present in different measure in every economic exchange.[1] It is well known that when non-contractible tasks are important, giving agents high powered incentives on contractible tasks or letting them compete on them - say, in a low price auction - may lead to a very inefficient outcome, for the principal and in general.[2] When exchanges are often repeated, underprovision on non-contractible dimensions can be avoided by governing them with reputational/relational forces.[3] But cooperation incentives based on "the shadow of the future" are typically stronger the higher are future expected payoffs at stake, so that enforcement may conflict with other important needs of a principal, and in particular with that of letting agents compete to screen and select the more able ones. This paper studies this conflict in a model where a principal recurrently trades with multiple heterogeneous privately informed agents, and where non-contractible dimensions may be important. We characterize the optimal relational contract defined in the broad sense to include equilibrium choices on explicitly contracted features and on the competitive screening policy – which are highly interdependent – identifying situations in which a tension between screening and enforcement indeed arises, and deriving the general implications of this trade off.

The importance of this tension is evident if we think, for example, at the difference between Anglo-Saxon and Japanese traditions in the management of suppliers and employees. The interaction between relational contracts and competitive screening is at the core of the comparison between Toyota's 'relational' procurement policies and, say, GM more 'competitive' or 'arms' length' ones (Asanuma, 1989). Toyota, as many other Japanese firms, maintains a small stable set of 'highly trusted' dedicated suppliers, restricts competition for the various orders only to them, caring for their profitability and rewarding the best performing suppliers with a higher share of orders, while replacing those that fail to deliver the extremely high levels of contractible and non-contractible quality required. The limits to competitive screening have a cost in terms of reduced screening and higher prices, at least compared to GM's more competitive selection of suppliers; but they ensure a sufficient weight to future stakes and a consequent cooperative perspective in the supply relation-

[1] Reasons why some dimensions of exchanges are not explicitly contractible include complexity and prohibitive legal cost of verification; see Hart (1995) for an in depth discussion and Tirole (1999) for an evaluation of the debate on contracts incompleteness.

[2] See Holmstrom and Milgrom (1991) and Manelli and Vincent (1995), among others.

[3] Macaulay (1963) is the classic reference.

ship. Analogously, a peculiar long-term perspective characterize Japanese employment relations: to ensure internal cohesion and a long-term 'loyal' orientation of employees, Japanese firms have for a long-time strongly limited competitive screening and other high-powered incentives in personnel policies, at least compared to their US counterparts (e.g. Aoki and Dore, 1994).

Public procurement is another obvious example. Banfield (1975) drew attention early on the likely loss of procurement quality caused by accountability rules that force public buyers to use open auctions for supplier selection. Kelman (1990) followed up on the theme, and on the importance of taking past performance and 'the shade of the future' into proper account. He stressed the differences between public and private procurement processes, noting that private firms use much less often open auctions, leave higher margins to suppliers, switch suppliers less often and are more satisfied about the quality of procured goods and service. When he became responsible for public procurement under the Clinton administration, he substantially reduced the rigidity of public procurement regulation to increase quality and reduce transaction costs, pushing for the collection of past performance information on all federal suppliers and for its use for their qualification in future tenders, so that relational-reputational forces could start working again. Europe has been instead moving in the opposite direction, pushing for open competitive screening whenever possible in the attempt to minimize discrimination of foreign suppliers and foster European market integration. This move, made consistent with a series of EU Procurement Directives, has forced many European countries in recent years to change their procurement regimes towards a generalized use of rigid and open competitive procedures above a certain threshold. Consistently with our results, a common feeling generated by these changes is that their main consequence has been a reduction in the quality of procured goods and services (besides an increase in transaction costs; see Europe Economics, 2006).

Outline of the paper. We develop and study a dynamic model of recurrent exchange with non-contractible tasks between an infinitely lived principal and a population of heterogeneous and privately informed infinitely lived agents. We therefore contemplate both moral hazard (on non-contractible dimensions) and adverse selection (on agents' type). In our environment, when the principal designs the optimal relational contract he must choose: any explicit part of the contract (enforceable by the court, like fixed payments/prices, contract duration, contractible performance standards); the implicit and self-enforcing incentives directed at governing the provision of costly non-contractible effort or quality; and the screening policy, that is if, how often, and how to let agents compete to be selected as suppliers.[4] These three aspects of the relationship are highly interdependent, so finding the optimal relational contract requires joint optimization on all three dimensions.

We first characterize the optimal relational contract and screening policy allowing parties to exchange contractible monetary transfers (wages, participation fees, payments related to auctions' bids) as well as non-contractible ones (like performance bonds and bonuses). The principal's temptation to renege on a promised bonus (or to withhold a bond) and then excluding the performing agent severely limits its use. Because of this, the buyer's most effective instruments to increase agents stakes and enforce non-contractible performance are (i) fixed transfers (similar to wages) regularly paid to all loyal agents the principal pre-select and admit in the pool of potential suppliers, with a cost increasing in the number of admitted agents; and (ii) expected future informational rents from limiting agent competition (at the cost of lower efficiency).

We find that when suppliers compete and both unconditional transfers and performance bonuses are available, the buyer optimally chooses:

(i) discretionary performance bonuses and open auctions, when non contractible dimensions are unimportant or there are few agents;

(ii) negotiation with a single agent on an indefinitely renewed contract conditional on the provision of non-contractible performance when this is crucial;

(iii) and in all other cases 'restricted procurement', that is, recurrent competitive screening among a stable subset of 'qualified' loyal agents (the smaller the more important non-contractible performance), under the threat of exclusion and replacement in case of low performance.

In several important situations direct negotiations and discretionary transfers, whether ex ante or ex post, are not available. This is often the case for large and bureaucratic firms or public organizations, where discretionary transfers are banned for accountability reasons. Very often these same accountability rules require to obtain competitive offers from a minimum number of potential suppliers in procurements (this is for example the case of most countries' small scale public procurement, and for international organizations like the United Nations). This leaves regulated procurers with the third option described above, and with that of open auctions without performance bonuses. We find that in this case procurers should optimally use restricted tendering under the threat of exclusion from the future set of invited bidders after low performance in almost

[4]Because we consider both contractible and non-contractible tasks, our is effectively a dynamic model with multi tasking in the sense of Holmstrom and Milgrom (1991).

all cases.

We also show that whenever non-contractible performance is valuable, it is optimal to shorten the duration of the (explicit part of the) contract. Abstracting from technological aspects such as the rate of obsolescence, a shorter explicit contract implies higher frequency of interaction, which makes it easier for the principal to obtain high non-contractible performance by threatening to withhold the bonus or exclude non-performing agents. Indeed, with more frequent contracting the threat of exclusion is closer in time and gains from "cheating" are smaller, so that larger implementable quality can be expected.[5] With competitive screening (open or restricted) shortening contract length also increases the average efficiency of the performing agents. At the same time, these desirable effects for the principal have to be contrasted with the costs of organizing more frequent auctions.

However, an environment with few and frequently interacting agents is also the most favorable one for inducing and sustaining cooperative behavior between the agents. We therefore illustrate and additional and rather general trade-off between reputational forces and collusive behavior. Longer duration of explicit contracts – less frequent screening – together with a larger pool of competing agents, deters collusion but also reduces non-contractible performance obtainable from the competing agents. Symmetrically, shorter contract duration and a smaller pool of eligible agents facilitate collusion but also the enforcement of non-contractible performance.

This new trade-off may appear disappointing because it seems to mark limits on the remedies that can be put at work for non-contractible tasks. However, our analysis clarifies that collusion itself can directly interact with agents' incentives to provide non-contractible performance. In fact, by increasing price, collusion clearly increases agents' gains from future trade with the principal which, as usual, can also be seen as the cost of being excluded when the principal reacts to low levels of non-contractible quality with exclusion. Hence, we show that there are conditions in which the trade-off between non-contractible performance and collusion is only apparent. For example, when non-contractible dimensions are very important and the principal cannot use participation fees (or there are few agents), the principal finds it optimal to induce agents to collude, and the

[5] Some recent studies confirm this intuition even in a procurement set up with important long term investment needs. Studying a data set for train operating companies in UK, Affuso and Newbery (2002) show that (discretionary) investment is stimulated by shorter rather than longer contracts. Notwithstanding a standard hold-up problem associated with contract renewal that should point in opposite direction, the authors suggest that frequent re-procurement with short contracts disciplines suppliers who care for future re-award of the franchise.

outcome with colluding agents is more efficient from a social point of view.

With a procurement interpretation this result appears somehow provocative as it suggests that a buyer may not be necessarily concerned by suppliers' collusion. However, what we explain is why cooperation among suppliers, very common in procurement in the forms of consortia, joint-ventures and other joint bidding agreements, is indeed likely in some circumstances to provide larger levels of quality and leave the buyer better off.

Related literature. Our paper contributes to the literature on efficiency wages and relational contracts that, from the pioneering contributions of Shapiro and Stiglitz (1984), Bull (1987), MacLeod and Malcomson (1989) and Baker, Gibbons and Murphy (1994) to the recent work of MacLeod (2003), Levin (2003) and Fuchs (2007), has formally analyzed the optimal design and the consequences of self-enforcing agreements for the governance of non-contractible tasks (see MacLeod 2007 for a survey).

The closest to our work is probably Levin (2003), who elegantly characterizes the optimal relational contract with moral hazard and adverse selection between a buyer and a single seller. There are several major differences with respect to our analysis. First, in our environment we allow for competition and several potential agents (suppliers) the principal (seller) can decide to contract with. A second related difference is that we focus on the interaction between various dimensions of the relationships: the explicit contractible ones, the implicit threats and rewards, and the competition and screening of agents. Other differences are that we look at the optimal relational contract from the point of view of the principal, who is the short side of the market, but that we don't introduce imperfect information on observable performance indicators. Our work is also close to MacLeod and Malcomson (1998), where relational contracts are studied between a number of principals and a larger number of competing agents, though in the absence of adverse selection and screening.

The possibility that a relational contract contemplates one principal with several privately informed agents that repeatedly compete for a task lasting a limited time span has received limited attention in the literature: Levin (2002) studies team production in relational environment but with no reference to competition or collusion; Rayo (2007) studies relational contracts between multiple agents, endogenously deriving organizational structure, but in a framework characterized by a team production problem with no adverse selection nor competition or collusion. A recently circulated paper by Board (2008) obtains a result analogous to (the second part of) our first result on the optimality of limiting the number of trading partners, but in a very different model where the principal is fully informed (so no competitive screening is needed or studied), and neither contractible monetary transfers (wages, prices) nor non-contractible ones (performance bonds or bonuses) are allowed for. It is therefore not entirely clear how this new paper relates to our result, nor whether there are real world situations relevant to both.[6]

Finally, our analysis also relates to the literature on reputation and competition, starting with the seminal work of Klein and Leffler (1981), Shapiro (1983), and Allen (1984). These early analyses were concerned with the compatibility of "quality-assuring" reputational equilibria – requiring rents that make the effort of maintaining reputation worthwhile also with free entry in the market – but did not analyze in detail firms' competitive interaction (firms' incentives to steal business from each other). Stiglitz (1989) raised the question how could reputation be compatible with perfect competition that should eliminate any future supracompetitive gains. Hoerner (2002) offers the first elegant answer to Stiglitz's question: in his model with heterogeneous consumers, adverse selection and moral hazard, high prices signal high quality and make competition compatible with (in fact necessary for) reputation to work. In our environment this solution would unfortunately not work, because prices are set in a competitive winner-take-all auction where a low price determines the winner or, in negotiations, are set by the buyer.

The rest of the paper is organized as follows. Section 2 presents the model setup. Section 3 provides analyzes procurement with competing agents, restricted competition and bilateral negotiation when the principal can use a full fledged set of instruments or she is limited by existing rules. Section 4 discusses the effect of agents' cooperation and collusion on implementable quality. Section 5 discusses some extensions of the base model and concludes.

2 The base model

At any period $t = 0, 1, 2, ...$ a principal (buyer) needs to procure a good or a task from one out of $N > 1$ agents (firms). The principal's per-period value of procurement net of payments $v(q_t)$ is increasing in a decision $q_t (\geq 0)$, the quality or effort, taken by the agent currently supplying in t, i.e. the *supplier.* The per-period cost for supplier i is $\theta_{it} + \psi(q_t)$ with $\psi(q_t)$ increasing, differentiable,

[6] In employment and procurement relationships, our main focus, both contractible and non-contractible monetary transfers are typically available.

concave and $\psi'(0) > 0$, $\psi(0) = 0$. The value of trade in period t with supplier i is $s(q_t) - \theta_{it}$ where $s(q_t) \equiv v(q_t) - \psi(q_t)$. Selecting adequate measure for q_t and appropriately scaling ψ, without loss of generality we set $v(q_t) = vq_t + v_0$ with $v \geq 0$ and $v_0 \geq 0$. Time horizon is infinite, 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, thus being not-contractible. For example, q_t may be the quality feature of a procured service that cannot be specified. Alternatively, it can be an effort (or specific investment) provided by an expert or an employee.

In our model, procurement can be seen as a multi-tasking activity which, in addition to q_t, contemplates a contractible decision taken by the supplier at cost θ_{it} that is worth v_0 to the principal and generates value $v_0 - \theta_{it}$. To avoid uninteresting cases we assume $v_0 > \theta_{it}$ for any θ_{it} so that the principal never wants to discontinue procurement.[7]

The principal does not know the cost θ_{it} of any agent i at any date t. Hence, she uses a (possibly) competitive search, the *auction,* that awards at a price b_w a procurement *contract* requiring the winning agent to supply for the next $x \geq 1$ periods.[8] Any standard auction format would do for our analysis, and to fix ideas we will explicitly refer to second price auctions. The cost of organizing an auction for the principal is $k \geq 0$. The principal may want to restrict participation to the auction, in which case she sets up a *pool of* $n \leq N$ *competing agents* and may even restrict n to 1 thus making take-it-or-leave-it offers to a single agent (a "degenerate" auction). Since they are ex-ante identical, if $n < N$ then agents admitted to compete are chosen randomly.

The principal and the agents may be able exchange other monetary transfers in addition to price for the procurement contract b_w. In particular, before competition takes place and agents learn their costs they may exchange *an ex-ante transfer* $w \in \Re$ (e.g. a fee $w > 0$ paid at any auction to participate in the selection process or a fix wage $w < 0$ to compensate agents for remaining available to the principal), and, at the end of any procurement contract, they may also exchange an *ex-post transfers* $B \in \Re$ (e.g. a discretionary bonus the principal may decide to pay or a bond posted by the supplier at the beginning of the contract that the principal may decide to retain discretionally). In the analysis we will also consider the possibility that some of these transfers are not feasible, for example, for institutional reasons.

[7] Our aim is to study simultaneous interaction between screening and enforcement, hence we will not consider multiple sourcing.

[8] To avoid integers issues, we will treat x as a continuous variable.

Since we want to leave any incumbency advantage out of the analysis, $\theta_{it} \in \Theta \equiv [\underline{\theta}, \overline{\theta}]$ is independently and identically distributed with density $f(\theta_{it})(> 0)$ and $\theta_e \equiv E(\theta)$ (see Section 5 for a discussion).[9] To simplify exposition we assume agents are fully informed (but the principal is uninformed). This may be justified in environments frequently interacting agents that know each other well (quite common in procurement, for example). This assumption is immaterial except for results in Section 4 in which case we will illustrate that the trade-offs underlying our results qualitatively persist also with asymmetrically informed agents.

Although quality is not contractible, the principal and her suppliers may still profit of ongoing interactions and reach some agreement, the *relational contract* C, about how they are going to behave in the future on and off equilibrium. C is self-enforcing if it describes a perfect public equilibrium of the repeated game (Fudenberg, Levine and Maskin, 1994). In the following we refer the explicit procurement contract lasting x periods simply as the "*contract*" and distinguish it from the implicit (or self enforcing) relational contract C. In our environment any equilibrium allocation supported by non-stationary strategies can be also supported by stationary strategies since incentives can be provided to all parties through immediate compensation without modifying future continuation payoffs (MacLeod and Malcomson, 1989 and Levin, 2003). Hence, we will concentrate on stationary contracts where behavior does not change on the equilibrium path of play.

The principal thus sets a relational contracts that, in addition to the contractible terms (e.g. the selection rules, contract length x and participation fee w), also specifies a required quality q that the supplier is expected to deliver in all x periods (that q is the same for x periods is without loss), the maximum number $n \leq N$ of competing agents and possibly also the discretionary bonus B that the principal may want to pay at the end of the contract.[10] Upon a deviation at date t, either by the date-t supplier or by the principal, C prescribes how parties will behave in the future. Following MacLeod and Malcomson (1989, 1998), the two separate forever whenever a deviation takes place (the most severe punishment, Abreu, 1988) and C prescribes that if the relationship between agent i and the principal breaks apart, whenever possible the latter replaces i with another agent from the pool of $N - n$ previously excluded agents, if $n < N$. However, with monetary transfers one can

[9] The assumption of IID costs is also in Levin (2003). That cost is drawn anew also in any of the x periods of a given procurement contract is immaterial for all results.

[10] Although not always realistic, assuming n contractible would strengthen our results. As in MacLeod and Malcomson (1989), bonds can be seen as negative bonuses. Bonuses and bonds exchanged between the principal and any agent that is not the current supplier are irrelevant.

always determine a self-enforcing relational contract $\tilde{C}$ supporting the same equilibrium payoffs as C but that does not require separation and is renegotiation proof (Levin, 2003).[11]

If the relationship between the principal and a supplier breaks apart, none of the other players knows the identity of the deviator. Consistently and as in most papers on relational contracts, we rule out external reputation effects that we will discuss in Section 5.[12]

The **timing of the game** is as follows.

$\mathbf{t} = -\mathbf{1}$: *The principal sets the relational contract C.*

$\mathbf{t} = \mathbf{0}$: *$n \leq N$ agents are randomly chosen and decide whether to participate in the infinite repetition of the following **"auction" stage-game**:*

– *At time $t_1 \geq 0$ the auction identify a unique supplier and the principal incurs the cost k;*

– *At any period $t \in \{t_1, ..., t_1 + x - 1\}$ the supplier procures q;*

– *At time $t_1 + x$, if nobody has cheated a new stage game begins, otherwise the principal decides whether to exclude or not the current supplier. Then, a new stage game begins.*

At date t the cost for any agent i to procure for x periods will be $\theta_{it} + \theta(x) + c(q, x)$ with $\theta(x) \equiv \theta_e(\delta - \delta^x)/(1 - \delta)$, and a cost of quality $c(q, x) \equiv \psi(q)(1 - \delta^x)/(1 - \delta)$. We define $\theta'(n)$ as the smallest element in $(\theta_{1t}, ..., \theta_{nt})$ and, to simplify notation, we will proceed as if the event $\theta_{it} = \theta_{jt}$ for some $i \neq j$ has zero measure at any t.

Before proceeding we conclude this section by briefly illustrating the benchmark of *"zero-quality" equilibria* in which at any t the supplier provides $q_t = 0$ and the principal requires $q = 0$ (and clearly sets $B = 0$) getting a payoff

$$V_0 = v_0 \frac{1}{1 - \delta} - [nw + E(b_w) + k] \frac{1}{1 - \delta^x}$$

In any auction with $n > 1$ competing agents, the most efficient one wins, gets a price $b_w = \theta''(n) + \theta(x)$ where $\theta''(n)$ is the (first period) cost of the second most efficient agent, and a profit $w + b_w - \theta'(n) - \theta(x)$. Hence, at any auction an agent may expect to earn a *profit* $w + \beta(n)\pi(n)$ where $\pi(n)$ is agent' expected *informational rent* $\pi(n) \equiv E[\theta''(n) - \theta'(n)]$ and $\beta(n)$ is the probability

[11] With $\tilde{C}$, upon a deviation any bilateral relationship is kept ongoing but the bilateral surplus produced by the date-t supplier and the principal is reallocated with appropriate transfers. The cheating party is punished to the outside-option and the other party obtains all the remaining bilateral surplus. This "Pareto-perfect" continuation is subgame perfect and guarantees that no surplus is wasted.

[12] However, we allow that if any of the n firms is directly harmed (i.e. getting a smaller payoff) by a deviation of the principal, both when i is or is not the current supplier, then the deviation is detected. One notable exception on external reputation is Levin (2002).

of being the most efficient one among the n agents.[13] The principal extracts $\pi(n)$ by optimally setting w which is thus a participation fee that agents in the pool have to pay. Since the expected cost $E(\theta'(n)) + \theta(x)$ of procuring a contract lasting x periods is decreasing in n, the principal prefers *open tendering* inviting all agents to bid, i.e. $n_0 = N$. Concerning contract duration, the principal then faces a simple trade-off. She compares the cost k of organizing a new auction with the associated price b_w reduction, i.e. the difference between the expected cost θ_e for one more period within the same contract and the cost $E[\theta'(n)]$ if instead she terminates the current contract and runs a new auction that selects the most efficient agent for the next period.

Lemma 1 (Zero-quality equilibria) *The optimal relational contract C_0 with $q = 0$ prescribes $B_0 = 0$, $n_0 = N$ and $x_0 = 1$ if $\theta_e - E\left[\theta'(N)\right] \geq k$, $x_0 = \infty$ otherwise.*

Notice that, since the cost saving $\theta_e - E[\theta'(n)]$ obtained with an auction with n agents is increasing in n, stronger competition induces (weakly) shorter contracts: n_0 and x_0 are substitutes. Furthermore, the principal may select a single seller and keep it forever so as to save the cost k of organizing auctions. However, it is immediate to see that this decision is always dominated by open tendering. We will indicate with V_0^* the principal's maximal payoff in a zero-quality equilibrium. Also notice that if the participation fee w were not feasible, the principal could not extract the informational rent $\pi(n)$ but the only difference would be that the actual first period cost for her would be $E[\theta''(n)]$ instead of $E[\theta'(n)]$.

Finally, were quality fully contractible, the optimal contract would reach the *first best* with the same contract length x_{FB} and number of competing agents n_{FB} as in C_0 but a quality q_{FB} that maximizes $v(q) - \psi(q)$.

3 Non-contractible quality, competition and negotiation

In any relational contract C_n with n competing agents that prescribes a strictly positive quality q, agents may be tempted to cheat and save on quality costs with two types of deviations. First, the supplier that has won an auction may decide to deliver zero quality along the current contract. Second, anticipating to cheat on quality in case of win and thus expecting lower costs with respect

[13] In case $n = 1$, we would have $\pi(1) = r - \theta_e - \theta(x)$ where r is the highest admissible bid. In this case, since the principal gives up screening, w and r are equivalent transfers and, to limit notation, we will only explicitly consider w. Hence, firm i's expected profit for a given contract is simply $\pi(1) = w - \theta_e - \theta(x)$.

to competitors, any agent may bid aggressively and win even if it is not the most efficient one. The principal can control both types of behavior by guaranteeing participating agents future profits larger than what they can get by cheating. In any future auction, any of the n agents may expect to benefit (i) from simple participation through the transfer w, and (ii) from winning the auction thus getting the (expected) price $E(b_w)$ and possibly receiving the bonus B. On the other hand, in case of win any agent expects a procurement cost $E(\theta'(n)) + \theta(x) + c(q, x)$, so that incentives to participate and deliver quality are characterized as follows.

Lemma 2 (Firms' incentives) *Consider a relational contract C_n prescribing $q > 0$. It induces the n agents to participate, bid planning to deliver q and actually deliver q if*

$$w + \beta(n)\pi(n) + B\ \mathcal{I}_{n=1} \geq 0, \tag{1}$$

$$[w + \beta(n)\pi(n) + B\ \mathcal{I}_{n=1}]\,\frac{\delta^x}{1-\delta^x} \geq c(q, x) - B. \tag{2}$$

Constraint (1) guarantees agents non negative expected profits at any auction. When competition is in place (i.e. $n > 1$), agents are forced to reduce their bid by any fix transfer that the principal may want to pay exclusively to the winner, in addition to b_w. Hence, the expected profit $w + \beta(n)\pi(n)$ ultimately does not depend on B, unless the principal prefers avoiding any form of competition and contracts with a single agent setting $n = 1$.[14] If agent' incentive compatibility constraint (2) is satisfied, any supplier that has won any given auction prefers not to cheat in the current contract since the present-value of expected future profits is larger than the cost $c(q, x)$ it can save by cheating on quality but then losing both the bonus B at the end of the contract and all future profits. Furthermore, any agent i at any auction never plans to cheat even if by doing so it can reduce its bid leading to a certain win at the auction. Indeed, with this deviation no agent (independently of its efficiency θ_{it}) can gain more than what is in the r.h.s. of (2) minus $\theta_{it} - \theta''(n) \geq 0$.

Condition (2) shows that both a larger w and B make cheating more costly for agents. Furthermore, a larger number of competitors n reduces both the probability of winning $\beta(n)$ and the informational rent $\pi(n)$. Similarly, a longer contract, i.e. larger x, makes rents more distant in the future and increases the cost of the current contract. Hence, stronger competition and longer contracts both have adverse effects on agents' incentives to deliver quality.

[14] In this case, however, we will show that the only credible bonus is $B = 0$.

If the bonus can be credibly employed by the principal, it is a very effective instrument to control agents' incentives since it can be retained at the end of the contract, whilst the fix fee w is paid in advance. This desirable property of B has been illustrated by several papers on relational contracting (e.g. MacLeod and Malcomson, 1989 and Baker et al., 1994). In addition, our environment illustrates some new and interesting features of the bonus that are specific to relational contracting with competing agents. First, as explained, the bonus is never actually paid an thus costs nothing to the principal since competition "sweeps it out" reducing the payment b_w precisely by the amount of B (as it would be to any fix transfer paid exclusively to the supplier in addition to b_w). Second, inducing incentives with a larger B may let the principal reduce the fix fee w that is instead paid to *all* the n bidding agents.

The possibility to use the bonus, however, may be limited by principal's incentives. To see this, consider the principal's expected present value payoff[15]

$$V_n \equiv s(q)\frac{1}{1-\delta} - [nw + E(b_w) + B + k\mathcal{I}_{n>1}]\frac{1}{1-\delta^x}.$$

At the end of any contract (having already paid b_w and w), the principal may renege the bonus. In this case, the relationship with the current supplier breaks down and the principal is left with her outside option $\underline{V}_n$, i.e. the expected payoff when the current supplier cannot be employed anymore.

It is useful to distinguish two procurement modes in terms of competing agents.

Definition 1 *Procurement is* open *if the principal sets $n = N$. It is* restricted *if $n < N$ and, in particular, it contemplates* negotiation *with a single agent if $n = 1$.*

We can then state the following.

Lemma 3 (Buyer's incentives) *Any relational contract C_n prescribing quality $q > 0$ is incentive compatible for the principal if*

$$(V_n - \underline{V}_n)\delta \geq B \tag{3}$$

and individually rational if

$$V_n \geq V_0^* \tag{4}$$

With restricted procurement, $B = 0$ in any principal-incentive compatible C_n. With open procurement there exist relational contracts C_n that satisfy (3) with $B > 0$.

[15]Notice that if $n = 1$ the principal needs not select the supplier so that she does not pay the cost k of organizing the search process.

When the principal sets $n < N$, thus inducing restricted procurement, she can always renege the bonus and replace the current supplier at no cost with another agent among those $N - n$ previously excluded from the pool of competitors. This guarantees a continuation payoff $\underline{V}_n$ that is at least V_n so that the unique credible bonus is zero.[16] If instead the principal admits all potential suppliers with open procurement, i.e. $n = N$, then a positive bonus may be credible. In fact, starting with all N agent, if now the principal reneges the bonus, she cannot replace the current supplier and her outside option $\underline{V}_N$ is determined by the optimal relational contract she designs with only $N-1$ agents, i.e. V_{N-1}.[17]. Finally, the principal's participation constraint (4) illustrates that she can always secure her zero-quality payoff.

We can now characterize the principal's optimal contract C_n^* that maximizes V_n subject to incentive compatibility and participation constraints of all players, i.e. (2)-(4). Since V_n is decreasing in the fee, w is optimally set such that (2) and (1) are binding, i.e.

$$w = \frac{1-\delta^x}{\delta^x} \max\{c(q,x) - B, 0\} - \beta(n)\pi(n) - B\ \mathcal{I}_{n=1}. \tag{5}$$

This shows that whenever the principal can control agents' incentives on quality by setting B equal to the cost $c(q,x)$, then w allows her to extract all the expected rents of bidding agents, as in the zero-quality equilibrium. In this case, w is a participation fee. If instead $B < c(q,x)$, then w may become negative thus being a transfer from the principal to the competing agents control their expected profits and incentives. In this case, the transfer w can be seen as an employment contract that the principal pay in order to have the n agents always available to perform procurement.

Substituting w and $E(b_w)$, the principal equivalently maximizes

$$V_n = s(q)\frac{1}{1-\delta} - [E(\theta'(n)) + \theta_e(x) + k\ \mathcal{I}_{n>1}]\frac{1}{1-\delta^x} - \frac{n}{\delta^x}\max\{c(q,x) - B, 0\} \tag{6}$$

subject to (3) and (4). This expression for V_n shows that, unless the bonus B exactly matches the cost $c(q,x)$, quality is distorted for incentive reasons with a distortion that increases with the number of competing agents n (the last term in (6)). The issue for the principal is that if $B < c(q,x)$

[16] This observation is related to efficiency wage-equilibria in the labor market where workers (in our model agents) are on the "long side" and firms (in our model the principal) are on the "short side" of the market (see for example MacLeod and Malcomson, 1998). In these cases as well bonuses are not a credible payments.

[17] As explained in the proof, upon a deviation, none of the $N-1$ remaining agents are negatively affected by the principal's deviation. Furthermore, after a deviation from a contract C_N, the possibility to set $n < N-1$ is dominated for the principal by contracting instead with all $N-1$ agents.

then, in order to give the right incentives to any prospective supplier to procure quality, she must increase the transfer w by the first term in the r.h.s. of (5) and since w must be paid to all competing agents (see the discussion above), the cost of giving incentives for quality is finally the last term in V_n.

Proposition 1 (Restricted and open tendering) *(i) The optimal contract C_n^* with* ***restricted procurement*** *is such that: q_n^*, x_n^* and n^* are downward distorted with respect to the first best; More important quality for the principal (i.e. larger v) induces a larger q_n^* but smaller x_n^* and n^*; Substitutability between n^* and x_n^* is stronger than in first best; If negotiation is optimal (i.e. $n^* = 1$) then $x_1^* = 1$.*

(ii) With ***open procurement****, there exists a decreasing function $\underline{N}(v) \geq 0$, with $\underline{N}(0) = \infty$ and $\underline{N}(v) = 0$ for any $v \geq \bar{v}(> 0)$, such that if the number of all agents is sufficiently small (i.e. $N \leq \underline{N}(v)$) then the relational contract is fully efficient (i.e. $C_N^* = C_{FB}$), otherwise q_N^* and x_N^* are downward distorted with respect to the first best and decreasing in N with $\lim_{N\to\infty} q_N^* = 0$.*

With restricted procurement the relational contracts always contemplates distortions since to provide incentives for the competing agents, the principal must reduce the level of quality she asks to the supplier, as indicated by the last term in (6). How much the principal wants to restrict competition depends on a trade-off on n. On one hand, a larger n reduces the cost of procurement via a smaller expected cost $E[\theta'(n)]$. On the other hand, a larger n reduces agents' expected rents thus making the incentive compatibility constraint of the supplier more difficult and costly to satisfy (as shown in (2) and in the last term of (6)). This trade-off may also lead the principal to set $n = 1$ thus, de facto, keeping renewing the contract with the same supplier (unless on of the two cheats) in which case it is optimal to set the shortest contract duration since with negotiation there are no costs k for organizing the search process. The optimal contract length for $n > 1$ instead, in addition to what illustrated for the zero-quality equilibria (Lemma 1), also accounts for the fact that a longer contract now increases procurement cost of quality $c(q, x)$ and again implies a larger incentive costs to control supplier's quality. The contract length is then shortened for incentive reasons. It is also worth noticing that if the principal wants to procure higher quality, then this increases the distortionary (last) term in V_n (i.e. $n/\delta^x\ c(q, x)$) and, consequently, the principal optimally further reduces both x and n.

Consider now open tendering. In this case the principal can credibly use the bonus and will in fact increase B as much as possible. What is more, and differently to standard relational contracting with a single agent, with competition the principal always optimally set B large enough such that the agent's incentive compatibility constraint (2) is always satisfied, i.e. raising B up to $c(q,x)$. Indeed, with competing agents a larger B makes incentive compatibility of the principal (3) "easier" to satisfy since the bonus is never actually paid and it allows to reduce the fee w paid to all N agents. As a consequence, the principal optimally sets $B = c(q,x)$ (an even larger B has no effect on V_N and is "costly" in terms of constraint (3)). This opens the door to the possibility that the optimal relational contract is fully efficient since the distortion in V_N vanishes. This possibility ultimately depends on the number of agents N. To see this notice that when quality and contract length are undistorted as in the first best, i.e. q^{FB} and x^{FB},, the difference $V_N - V_{N-1}$ in the l.h.s. of (3) simply reflects the (cost-)efficiency gain that the principal obtains with one more agent instead of contracting with $N-1$ agents, i.e. $E[\theta'(N-1)] - E[\theta'(N)]$) which is decreasing in N. Hence, if N is sufficiently small, then it may well be the case that $(V_N - V_{N-1})\delta$ is larger than $c(q^{FB}, x^{FB})$ and the relational contract is fully efficient. If instead N is large, the (present value of the) gain of having one more agent is smaller than the cost to procure the efficient level of quality. In this case, principal's incentive constraint (3) is not satisfied and the principal must reduce both q and x. Furthermore, since $V_N - V_{N-1}$ is bounded from above by $E[\theta'(N-1)] - E[\theta'(N)]$ which is decreasing in N, quality and contract duration are both decreasing in N and respectively tend to zero and $x = 1$ as N becomes larger and larger.

In light of Proposition 1, what is the best procurement mode after all? Clearly, if N is sufficiently small, open tendering may allow the principal to reach the first best and any other procurement mode is dominated. However, we know that when then number of competing suppliers is large enough, procurement with open tendering is distorted as well and the comparison between the two procurement modes depends on two different distortions.

Corollary 1 *There exists a decreasing function $\overline{N}(v) \geq \underline{N}(v)$, with $\overline{N}(v) - \underline{N}(v)$ increasing in v and $\overline{N}(0) = \underline{N}(0)$, such that for $N \leq \underline{N}(v)$ open procurement is optimal whilst for $N \geq \overline{N}(v)$ restricted procurement is optimal with a number of competing agents $n^*(v)$ which is (weakly) decreasing in v, in the limit $n^*(v) = 1$ (negotiation). In intermediate cases, i.e. $N \in (\underline{N}(v), \overline{N}(v))$ optimal procurement may be open or restricted.*

If quality is not at all important for the principal, i.e. $v = 0$, we already know that maximal competition among agents is the best choice for the principal, i.e. open procurement. This is also the case for larger but not too large v and the number N of competing agents is not very large since we know that in this case open procurement allows eliminate any distortion. If instead quality is very important for the principal, i.e. v is sufficiently large, then she has to control suppliers' incentives to deliver q by limiting competition. Indeed, open procurement is associated with a strong distortion since the bonus is very limited by the principal's incentive constraint and open competition guarantees too small rents to the agents. In this case then the principal prefers restricted procurement so as to control the number of competing agents n and their rents and this can also lead to negotiation with $n = 1$ for very high v.[18] Summarizing, the least distorted quality is obtained by the principal, either with open tendering if both N and v are small, or with restricted procurement and, in particular, negotiation, otherwise.

3.1 Procurement with limited instruments

As illustrated in the Introduction, the principal is often unable to use discretionary bonuses and / or participation fees w, this may be the case, for example, for some public buyers. It is then important to verify whether the previous results are robust to a restriction on the instruments the principal can use to govern procurement.

Corollary 2 (Procurement with no B and w.) *If the principal cannot use the bonus B and / or the fixed transfers w, then the optimal number of competing agents n is further reduced and restricted procurement is optimal for a larger set of parameters.*

In this case, implementable quality is, ceteris paribus, lower the less heterogeneous the agents are.

Suppose the principal cannot use the bonus. Hence, the only difference with restricted procurement is the larger number of competing agents with open procurement. If the principal cannot even use the fixed fee w either, then the only way to govern supplier's incentives to procure quality is through supplier's expected informational rent $\beta(n)\pi(n)$, because agent's incentive compatibility constraint (2) becomes

$$\beta(n)\pi(n) \geq \frac{1-\delta^x}{\delta^x} c(q,x) \tag{7}$$

for any n. Since the expected informational rent $\beta(n)\pi(n)$ is decreasing in n (both terms decrease with n), the only possibility to increase q in an incentive-compatible way is to further restrict the number of competing agents (with the highest implementable quality being obtained with

[18] Clearly, larger cost k of organizing auctions makes contract renewal relatively more desirable, ceteris paribus.

negotiation).[19] For future reference we notice that when w cannot be used, condition (7) implicitly defines the *maximal implementable quality* $q(n,x)$ decreasing in n and x, that the principal can ask for any n and x (defined as q that makes (7) binding) so that any q is incentive compatible if and only if $q \leq q(n,x)$.

When the set of instruments of the principal is limited and she must rely on the expected informational rent to govern quality, then a small variability of agent' cost θ reduces the level of implementable quality. In the limit, when agents are homogeneous and $\pi(n) = 0$ for any $n > 1$, the principal should procure with negotiation if quality is important for the principal since this is the only way to guarantee agents some future rents. With this respect it also worth noticing that some (public) buyers are not allowed to completely eliminate competition setting $n = 1$ (in some countries a public buyer must invite a minimum number of firms to present an offer) so that another limit in the set of procurement instruments may be the impossibility to rely on negotiation with a single agent.

4 Reputation and collusion

We have shown that controlling for non-contractible quality may require a reduction in contract length x (Proposition 1) and of the number of competing agents n (Corollaries 1, 2), when quality is important for the principal (i.e. v high). It is known that both these actions are (among) the most effective conditions that foster and strengthen collusion (i.e. cooperation, we will use the two terms interchangeably) among agents. Since collusion in procurement is far from being a simple theoretical curiosity and, on the contrary, it is a pervasive phenomenon, our results in the previous Section point out that, controlling for quality, the principal may risk to induce agents to collude. We now explicitly address this issue and illustrate what may happen to the procurement process if competition cannot be simply assumed.

To account for the possibility that agents collude, we now indicate with $\tilde{\pi}(n)$ the expected rent that a supplier can obtain if a cartel is in place among the n agents. In this case, the most efficient agent is awarded the contract and all the other potential suppliers in the pool of n agents either abstain from bidding or submit losing bids, i.e. collusion takes place with bid rotation. For the purposes of this section it is convenient to define the highest acceptable bid r set by the principal

[19] A similar reasoning applies if the principal can use B but not w, as we show in the Appendix.

so that the winning bid of colluding agents will be equal to r.[20]

We start letting the principal use the complete set of instruments and then discuss the effects of limited instruments as in Section 3.1. Since cooperation among agents affects their payoff and also their incentives to deviate, we now explicitly let the principal conceive a relational contract anticipating that agents will be induced to cooperate.

For cooperation to be sustainable at any auction, the second most efficient agent (i.e. the one with the highest incentive to cheat on cooperation) should not prefer to undercut the currently most efficient one. If it does not deviate, he can expect the future collusive profits. Alternatively, by deviating he obtains an immediate gain $D \geq 0$ but then cooperation breaks down and all agents will compete forever. Clearly, a deviating agent may also contemplate the possibility to cheat on quality. Hence, cooperation with provided quality q is viable if the following incentive compatibility constraint is verified,

$$[w + B + \beta(n)\tilde{\pi}(n)] \frac{\delta^x}{1-\delta^x} \geq D + \max\{B + [w^* + \beta(n)\pi^*(n)] \frac{\delta^{x^*}}{1-\delta^{x^*}}, c(q,x)\} \tag{8}$$

where the variables indicated with an asterisk refer to the optimal relational contract among competing agents derived in the previous section since, if cooperation brakes down, the principal realizes that agents start competing and accordingly adjusts the relational contract.[21]

Now, when cooperation is in place and it is stable (i.e. condition (8) is satisfied), the current supplier prefers not to cheat on quality if

$$[w + B + \beta(n)\tilde{\pi}(n)] \frac{\delta^x}{1-\delta^x} \geq c(q,x) - B \tag{9}$$

and notice that with cooperation the supplier does earn and the principal does pay the bonus B since cooperation avoids that competing agents reduce their bids by B.

Since the left hand side of the incentive compatibility constraint (8) of cooperation is decreasing in x and the right hand side is (weakly) increasing in x, then by reducing contract length x the incentives to cooperate become stronger. Similarly, a smaller n increases the probability $\beta(n)$ of being the most efficient agent and, since also $\tilde{\pi}(n) \geq \pi^*(n)$, a smaller pool of potential suppliers

[20] We model collusion in a simple way avoiding several complications. We do not consider partial collusion that involves less than n agents and assume that, if collusion is incentive compatible, it is so and takes place for any realization of costs. Alternatively, for low realization of costs the cartel may contemplate temporary reversion to competitive bids (as in Rotemberg and Saloner, 1986), but this would not qualitatively alter our results.

[21] The winning bid tells the principal whether cooperation has felt apart or not. Assuming that the relational contract remains unchanged after cooperation breaks down may fail to satisfy subgame perfection.

n makes cooperation stronger.[22] As anticipated, when the principal wants to increase the level of quality, at the same time she increases the scope for and stability of cooperation and may end up with the following trade-off: trying to implement larger quality with reputational forces, the principal increases the risk of inducing potential suppliers to collude.

These reasoning immediately lead to the following.

Proposition 2 *There is a general trade-off between reputation and collusion: setting smaller x and n to increase larger q via larger agents' future profits, the principal tends to induce collusion among agents.*

Interestingly, this is a rather general trade-off that is clearly relevant in many other environments that share the common ingredients of potential competition among agents and the need to give them the incentives to perform non-contractible tasks.

How does the principal address this trade-off between non-contractible quality and collusion? What is optimal procurement in light of this trade-off? The following Lemma provides a partial answer to these questions.

Lemma 4 (Quality with firms' cooperation) *If quality q is implementable with competing agents, it is certainly so with colluding agents.*

This Lemma unexpectedly shows that collusion among agents may be interesting for the principal, contrary to common wisdom. This is a simple consequence of the fact that the stability of cooperation (i.e. constraint (8)) clearly implies that the expected profit with cooperation is larger than with competition. Hence, any supplier when deciding whether or not to deliver q knows that the cost of cheating on quality is larger than with competition since the profits at stake are larger. It is important to clarify that, as stated in the Introduction, what we indicate as cooperation here can be seen and interpreted as a metaphor for any agreement such as collusion among agents but also self-sustaining consortia among potential suppliers and other forms of joint bidding like joint ventures. In this case, the principal would equivalently negotiate with a single consortium composed by the same number of agents and incentive compatibility could be seen as internal incentives for the stability of the consortium independently of any legal obligation among agreeing partners. It

[22] We omit for simplicity of exposition the fact that, for given collusive bid, a smaller n also reduces D since the second most efficient agent is more efficient and that $\pi(n) - \pi^*(n)$ is also decreasing in n. Both these effects strengthen the reasoning in the text.

is also worth mentioning that Lemma 4 also holds independently of the possibility for the principal to use B or w. Indeed, in case these instruments are not available, the maximal implementable quality is $q(n, x)$ (as defined by agents' incentive compatibility constraint (7)) which is larger the larger is the expected informational rent.

We can then ask whether the principal can be ultimately better off with procurement that induces agents to cooperate.

Proposition 3 (Procurement with collusion and consortia) *Assume the fee w is not a feasible instrument for the principal. For a value of quality v sufficiently large, then procurement with cooperating agents is optimal when either (a) the principal cannot procure with negotiation, or (b) she also cares for procurement efficiency in addition to V_n.*

If instead w is feasible, then procurement with cooperating agents is optimal only if N is not too large and $B > 0$.

Imagine first the principal cannot use the transfer w, as it is sometimes the case, as argued before. Proposition illustrates that when quality is important for the principal limiting competition may become desirable in addition to what stated in Proposition **??**. Not only the principal may optimally want to restrict the number of competing agents, she can further increase quality by inducing agents to cooperate within consortia or cartels since, inducing cooperation, agents' expected profits increase and this ultimately allows to implement larger quality. Furthermore, since B with colluding agents is more effective than with competing agents in controlling profits (B is indeed paid in equilibrium), the principal tends to prefer open procurement (i.e. $n = N$) with colluding agents more than with competing agents. In fact, imagine constraint (3) does not bind. Since B and w have now equivalent effects on the l.h.s. of (9), the principal can increase B and reduce w by the same amount thus guaranteeing the incentives for quality and reducing their cost (recall w must be paid to all n agents).

Clearly, if the principal can restrict procurement to negotiation, then this gives the largest possible rent to the single supplier. However, we know that there are cases in which a public buyer, for example, cannot completely eliminate competition so that inducing agents to coordinate turns out to be optimal when quality is sufficiently important.

Again in public procurement, a general concern of the buyer may also contemplate efficiency in production. If this the case, it is clear that the larger rent that could be given to the single supplier with negotiation must be confronted with the efficiency reduction of the single agent as compared with $n > 1$ cooperating agents. The balance between high quality and high efficiency ultimately induces the principal interested in high quality and efficiency either to chose negotiation (if admissible) or cartel / consortium formation.[23]

Consider now the case in which the principal can use the transfer w. The benefits from collusion are now reduced since she can control agents' profits and then quality directly with w. However, when we now compare competition with collusion then the fact that in the latter case B is indeed paid to the suppliers allows to accordingly reduce the transfer w paid instead to all the agents, as compared with competition. Hence, if $B > 0$ and N is not too large (otherwise we know from the analysis in the previous section that B must be low to control the principal's incentives), then the positive effect of B is sufficLiently large with collusion so as to make the principal prefer it to competition whenever competition with N agents was optimal in the previous Section.

The procurement literature has often stressed the benefits of consortia and joint bidding among otherwise rival firms. However, these agreements are almost systematically treated in the literature as black boxes. Our analysis instead allows to illustrate the benefits (here in terms of larger implementable quality) but also the possible limitations of these agreements by explicitly accounting for their stability requirement. Furthermore, in some cases consortia are ruled out by law and then inducing agents to implicitly cooperate is a way to bring back the potential benefits of cooperation in terms of larger implementable quality.[24] In other cases, consortia are admissible and do not need to be self enforcing (being governed by explicit contracts). If this is the case and cooperating among agents is optimal in terms of our previous analysis, a consortium is then even better since it guarantees larger rents to the firms and the same level of efficiency as with collusion.

5 Extensions and Concluding remarks

In the previous sections we have illustrated how a principal addresses the tension between screening and enforcement when heterogenous agents are available for procuring a good or a task of non contractible quality. The simple environment we have used is rich enough to deliver insights on the

[23] The cartel is here powerful in sorting out the most efficient agent. On the other hand, it is also clear that even if collusion were less efficient, still the potential of improvements in efficiency are larger as compared with a single agent and this efficiency enhancement is exactly the main driver in the formation of consortia.

[24] For example, in the case of public procurement, the Italian antitrust authority recently ruled out the possibility for consortia to participate, unless firms in the consortium are able to prove that they could not technically supply the buyer if they were obliged to operate independently.

level of competition and contract length that are optimal for the principal. In particular, we have shown that when the value of quality is important, the principal tends to prefer shorter contracts and limited competition that may also lead agents to cooperate (i.e. collude) in the interest of the principal. Although the model is simple, it may accommodate some extensions that are worth mentioning.

Information and renegotiation. Notwithstanding our assumption on fully informed agents, the drivers of our simple trade-offs also hold in a more complex environment with privately informed agents. Indeed, the analysis in Sections 2 and 3 can be extended to asymmetric information almost without modification (applying the revenue equivalence theorem all standard auction formats would remain equivalent for the principal).[25] As for the analysis of cooperating agents, significant complexities have been illustrated by the literature on collusion with repeated auctions (see Aoyagi, 2003 and Blume and Heidhues, 2004) and clearly being the cartel members privately informed, the efficiency properties of the cartel would be weakened. However, what matters for our results on collusion is the simple fact that cooperation is incentive compatible since, as shown in the previous section, this necessarily implies that (equilibrium) expected profits with collusion are larger than with competition so that, ultimately, the implementable quality is larger. Furthermore, the comparison between procurement with cooperating agents and that with a single agent is also qualitatively unaffected since it is clear that, on average, the efficiency of the supplier selected out of the many cooperating agents is larger than that of the single agent with negotiation.

In our analysis we deliberately avoided any incumbency advantage to agents contracting first with the principal since the aim of the analysis was the relationship between screening and non-contractible dimensions. Hence, for this reason we have assumed that agents' efficiency is IID (across periods or, equivalently, auctions), as it is assumed in all papers on relational contracting with asymmetric information, to our knowledge. Had we assumed cost persistence, the principal would have learned from auctions and the cost of dismissing a cheating but efficient agent would have been higher than in our environment. Set aside the complications of such a model, in this case exclusion could be less of a scarecrow for efficient agents thus pointing to an intrinsic trade-off between efficiency and quality. On one hand, less efficient agents would be conscious that they can be easily discarded and substituted and this provide the right incentives for quality provision.

[25] We did not consider the possibility to use "Maskin mechanisms" in which the principal would punish costs reports by agents that are not consistent since this possibility would not be available with privately informed agents.

On the other hand, more efficient agents know that the principal would be reluctant to discard them and are then less disciplined to provide high non contractible quality.[26] The analysis of this is novel and interesting trade-off requires to consider non-stationary relational contracts and is left for future work.

Cartels, joint bidding and sub-contracting. Cartels and joint bidding (or equivalently consortia) may act even more efficiently than what discussed in the previous sections. Indeed, the cartel can further boost efficiency delegating (or subcontracting) production to the agent that is period-by-period the most efficient one. In this case, clearly, the optimality of procuring with a cartel or a consortium increases with respect to both competition and single supplier. For simplicity we also have considered only consortia or cartels that comprise all the n agents admitted at the bidding stage but this is not the unique possibility. The principal may induce partial cartellization, for example allowing consortia of maximum two agents so that the incentives to deliver non contractible quality, the price to procure and the efficiency could be further "fine tuned" to the principal's interest.

Another ingredient we have not explicitly addressed is the possibility that a cartel or a consortium may also help agents to monitor principal's behavior. The idea is that the common organization may allow all cooperating agents (and not only the current supplier) to verify quality and then check principal's deviations. This possibility is known in the literature as multilateral relational contracting (see Levin, 2002), to be contrasted with the case of bilateral relational contracting analyzed in our paper. Since with multilateral relations the cartel or consortium can pool the reaction of many agents when the principal deviates, this may discipline the latter and also let implement larger non-contractible quality.

Finally, sub-contracting often takes place in procurement and its effects on screening and non contractible quality may turn out to be intricate. Indeed, responsibility for quality provision may remain in the hands of the main contractor so that incentives for quality may well result diluted. We plan to investigate this interesting relationship between sub-contracting and non-contractible quality in procurement in a future work.

[26] Although in our model changing efficiency during the contract execution seems to point to the possibility of beneficial renegotiation, it is worth emphasizing that the trade-off between efficiency and non-contractible quality is already accounted for by the principal (also) with the choice of contract duration. Hence, a fortiori with any specific cost for renegotiation this would not take place in the current setup.

Minimum price, scoring rules and alternative mechanisms. In the previous pages we have used a mix of transfer ex-ante and ex-post that allow the principal to control agents' future profits. With this respect the approach is general enough to deliver results that are robust to possibly more complex mechanisms. First, recall that in the current environment stationary relational contracts are without loss of generality. Furthermore, the instruments that are available to the principal allow her to control agents' profits and incentives to deliver non contractible quality and to screen on procurement costs. As explained above, there are no other types of transfers that the principal may use to increase the suppliers' profits (and then control quality) in addition to the fee w paid to all bidding agents, the price to the winner b_w and the bonus B.

We did not explicitly consider the possibility that the principal sets a minimum price l so that admissible bids must be $b_i \geq l$. Minimum prices are often used in procurement when a buyer cannot restrict and pre-select the potential suppliers, particularly in public procurement and, on the contrary, they are rarely employed in private procurement when the buyers can limit the number of eligible suppliers. Interestingly, our analysis sheds light on this observation. When agents cooperate the presence of a minimum price is irrelevant, as well as when a single supplier is admitted. When instead agents compete and the principal does not care for efficiency, a larger minimum price is substitute to a reduction in n. In fact, with a given probability the realizations of the costs of all agents are larger than the minimum price in which case nothing changes on agents' expected profits. With complementary probability, instead, the procurement cost of the most efficient agent is below the minimum price and two sub-cases may take place. First, the costs of all other agents are above the minimum price and nothing changes so that the most efficient agent wins the auction with the same profits as without the minimum price. Second, the cost of at least another agent is below the minimum price. Let consider the simple case in which this is true only for the second most efficient agent (a similar reasoning applies in general). Hence, both these agents bid the minimum price and have equal probability to win. The expected profit of the most efficient agent is now lower, and that of the second most efficient is larger, than without the minimum price in a way that, ex-ante, the expected profit of any agent is in fact unchanged.

Finally, the principal may also envisage the possibility to rank agents' offers according to a scoring rule which is a function of the bid (i.e. the price) and the quality the agent promises to deliver. Instead of relying on the simple quality as in the previous pages, the principal would then exclude a supplier that failed to generate ex-post the implicitly promised score.[27] However, this

[27] The possibility to use these scoring rules may be limited by the fact that the assignment of the contract (i.e. a con-

form of competition with bid-quality offers and scoring rule is irrelevant as long as there is no cost heterogeneity among firms for quality, as in our model.

Subjective quality evaluation. The recent theory of relational contracts with subjective performance measures (Levin 2003, MacLeod 2003 and Fuchs 2007) has emphasized often the realized quality observed by the principal may be subject to noise, so that the principal and the agent have private information on what they observe. With this respect, a common theme is that to induce the principal to report truthfully the perceived quality and act according to the prescriptions of the relational contract, the optimal contract must make the principal indifferent between reporting different performance levels of the agent. In relationships with a single agent this tends to induce inefficiencies (sometimes obtained with "money burning"): when the agent's performance is poor and the contract prescribes a punishment for the agent, the principal must not gain from that punishment, otherwise it would be induced to report bad performance of the agent more often.

In this paper we have illustrated that exclusion of some potential agents is indeed optimal in most relevant cases. In particular, we have shown that with restricted procurement the buyer does not gain from punishing a firm that did not perform, as all firms are identical and replacing one with another brings nothing to him. This means that if instead of the observable quality we would have assumed a subjective quality assessment from the buyer, we would have found analogous equilibria and related results. The reason why no additional inefficient "money burning" is required –already pointed out in Fuchs (2007)– is that, differently than in bilateral relationships, the presence of competing agents allows the principal to punish the incumbent for its poor performance without gaining anything from punishing, rather, having a competing agent to profit from it. This maintains incentives for truthful reporting by the buyer.

References

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

tractible dimension) turns out to be determined also by non-contractible dimensions. In case of public procurement, for example, this may not be viable.

[2] Affuso, L. and D. Newbery (2002), "The Impact of Structural and Contractual Arrangements on a Vertically Separated Railway", *The Economic and Social Review*, Vol. 33, 1: 83-92.

[3] Allen, F., (1984), "Reputation and Product Quality", *RAND Journal of Economics*, 15, 311-27.

[4] Aoyagi, M., (2003), "Bid Rotation and Collusion in Repeated Auctions", *Journal of Economic Theory*, 112, 79-105.

[5] Banfield, Eric, (1975), "Corruption as a Feature of Governmental Organization", *Journal of Law and Economics* 18, 587-605.

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

[7] Blume, A. and Heidhues, P. (2004), "Modeling Tacit Collusion in Auctions", mimeo University of Pittsburgh and WZB Berlin.

[8] Bull, C. (1987), "The Existence of Self-Enforcing Implicit Contracts", *Quarterly Journal of Economics*, 102(1), 147-159.

[9] Europe Economics (2006), *Evaluation of Public Procurement Directives*, European Commission, DG Internal Market.

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

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

[12] Hart, O., (1995), *Firms, Contracts and Financial Structure,* Clarendon Press, Oxford.

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

[14] Kelman, S., (1990), *Procurement and Public Management: The Fear of Discretion and the Quality of Goverment Performance,* AEI Press.

[15] Klein, B. and Leffler, K., (1981), "The Role of Market Forces in Assuring Contractual Performance", *Journal of Political Economy*, 89, 615-41.

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

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

[18] Macaulay, S., (1963), "Non Contractual Relations in Business: A Preliminary Study", *American Sociological Review*, 28, 55-67.

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

[20] MacLeod, W.B. (2007), "Reputations, Relationships, and Contract Enforcement," *Journal of Economic Literature*, vol. 45(3), pages 595-628.

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

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

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

[24] Milgrom, P. R. and Roberts D. J. (1982): "Predation, Reputation and Entry Deterrence," *Journal of Economic Theory*, 27, 280-312.

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

[26] Rotemberg, J. and Saloner, G., (1986), "A Super-game Theoretic Model of Business Cycles and Price Wars during Booms", *American Economic Review,* 76(3), 390-407.

[27] Shapiro, C., (1983), "Premiums for High Quality Products as Returns on Reputation", *Quarterly Journal of Economics*, 98, 659-80.

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

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

[30] Tirole, J., (1999). "Incomplete Contracts: Where Do We Stand?", *Econometrica*, 67(4), 741-782.

Appendix A

Proof of Lemma 1. The derivation of buyer's optimal procurement with $n > 1$ agents is immediate. We substitute w from the binding agents' participation constraint (**??**) into the buyer's objective function and, using $\beta(n) = \Pr[\theta_{it} = \theta'(N)] = 1/n$, we get $V_0 = v_0 \frac{1}{1-\delta} - \left\{E[\theta'(n)] + \theta_e(x) + k\right\} \frac{1}{1-\delta^x}$. The optimal C_0 simply follows from maximization of V_0. Finally notice that the fee w takes place before agent i learns its type θ_{it}, thus we are considering ex-ante participation for the agent. Allowing w to be paid after the agent learns its type would not alter qualitatively our results. In this case the principal may be able to extract some informational rents from the unique supplier, but still less than when allowing $n > 1$ agents to compete.[28] ■

Proof of Lemma 2. Consider any auction taking place at t. Clearly, if any agent prefers to bid in t anticipating it will deliver quality q in case of win, it will prefer to do so at any future auction in $t' > t$. Furthermore, we next show that in a SPA in which all agents plan not to cheat, truthful bidding is a (weakly) dominant strategy.[29]

Suppose agent i is the most efficient agent in t. When all other agents bid truthfully planning to deliver q, by truthful bidding and planning not to cheat on q this agents i wins the auction and obtains the expected payoff indicated in the l.h.s. of the following

$$b_w - \theta_e(x) - \theta'(n) - c(q,x) + \{w + \beta(n)\pi(n)\} \frac{\delta^x}{1-\delta^x} \geq b_w - \theta_e(x) - \theta'(n) - B$$

[28] We are not considering extremely large values of k that would make the principal prefer not to screen at all and contract with a single agent, i.e. setting $n = 1$. This possibility would make the analysis clearly not interesting both when $q = 0$ and $q > 0$.

[29] In the proof we will make use of the properties of the SPA and of full information among agents. However, it will be clear that none of the results really hinges on these assumptions.

where the r.h.s. is instead is payoff if he cheats on quality. It is also immediate that i has no incentive to bid untruthfully and / or planning to cheat on q. Hence, constraint (2) controls both ex-ante (at the bidding stage) and ex-post (at the quality setting stage) incentives for the most efficient agent.

We consider now the possible deviation by agent i who is not the most efficient in t and may bid planning to cheat. If agent i anticipates to deliver q in case it wins, it will bid $b_i = \theta_{it} + \theta_e(x) + c(q,x) - B$. On the contrary, planning to cheat on quality its costs would be $\theta_{it} + \theta_e(x)$. If he bids planning not to cheat, he will not win in t, still obtaining a payoff equal to w plus $\{w + \beta(n)\pi(n)\} \frac{\delta^x}{1-\delta^x}$, from future auctions. If he bids planning to cheat, let us consider the most favorable case to this agent, so that he bids the price of the most efficient agent (who were supposed to win in t) i.e. $b_w = \theta'(n) + \theta_e(x) + c(q,x) - B$ and he certainly wins. His payoff would be $w + b_w - \theta_{it} - \theta_e(x)$ and so that he will prefer to bid not planning to cheat if

$$\theta_{it} - \theta'(n) + \{w + \beta(n)\pi(n)\} \frac{\delta^x}{1-\delta^x} \geq c(q,x) - B \tag{10}$$

Since $\theta_{it} - \theta'(n) \geq 0$ it is immediate that (2) implies (10). ■

Proof of Lemma 3. Constraint (3) controls principal's incentives. If it is not satisfied the principal prefers retaining B and reverting the the outside option $\underline{V}_n$. Clearly, if $V_n = \underline{V}_n$ the principal faces no cost reneging B and it must be $B = 0$. This is the case when $n < N$ since the principal can substitute the current supplier with on in the $N - n$ pool of previously excluded agents. If instead $n = N$, then reneging B has the cost of losing one potential supplier so that $\underline{V}_N$ is the maximal payoff the principal can get with $N - 1$ agents and then clearly $V_N \geq \underline{V}_N$. Notice that after a deviation from a contract C_n with $n = N$ agents, the possibility to turn to a contract that contemplates $n < N - 1$ agents is dominated for the principal by contract with $N - 1$ agents. Indeed by so doing $(N - 1) - n$ agents would be negatively affected since excluded from the new relational contract and would not participate anymore in the future. Thus, the number of potential suppliers would be in fact n and not $N - 1$, with no possibility to replace a supplier in any case.

Furthermore, none of the $N - 1$ agents would be negatively affected by the change of the relational contract designed by the principal after her own deviation and reducing the number of agents from N to $N - 1$. Indeed, the optimal relational contract with open procurement and $N - 1$ would leave these agents with the same expected payoff.

As for participation constraint (4), starting with $n > 1$ agents the principal has always the possibility to give up with non contractible quality thus getting V_0^* or negotiating with just one agent with a payoff V_1^*. ■

Proof of Proposition 1. (i) Since $B=0$ the principal maximizes

$$V_n = s(q)\frac{1}{1-\delta} - [E(\theta'(n)) + \theta_e(x) + k\mathcal{I}_{n>1}]\frac{1}{1-\delta^x} - \frac{n}{\delta^x}c(q,x)$$

w.r.t. q, n and x.

The optimal q^*, n^* and x^* satisfy the following conditions

$$\frac{ds(q)}{dq} = \frac{n}{\delta^x}(1-\delta)\frac{dc(q,x)}{dq}, \tag{11}$$

$$k\mathcal{I}_{n>1} \geq \{\theta_e - E[\theta'(n)]\} + \frac{n}{\delta^x}c(q,x)\ \frac{1-\delta^x}{\delta^x}, \tag{12}$$

$$\Delta\theta(n-1)\frac{1}{1-\delta^x} \geq c(q,x)/\delta^x \geq \Delta\theta(n)\frac{1}{1-\delta^x}. \tag{13}$$

where $\Delta\theta(n) \equiv E[\theta'(n)] - E[\theta'(n+1)] \geq 0$ if $n>1$ and $\Delta\theta(n) \equiv \Delta > 0$ if $n=1$ with the scalar $\Delta > c(q,x)/\delta^x$ for any q and x.

Since the r.h.s. in (11) is positive and increasing in n and x, it follows that $q^* \leq q_{FB}$ and q implicitly defined by (11) is decreasing in x and n.

If (12) is satisfied with a strict inequality then $x^* = \infty$. Since, in the first best, the optimality condition for x would be the same but with last term in the r.h.s. (i.e. the distortion) equal to zero, it follows that $x^* \leq x_{FB}$. Notice also that if $n=1$ then (12) is impossible and $x^*=1$. Furthermore, $x=\infty$ is now never optimal since this would imply that (12) is never satisfied and the optimal x would be $x=1$, a contradiction, hence, $x^* \in [1,\infty)$. Notice also that a larger n increases the r.h.s. in (12) thus inducing a smaller x and this is true not only for the curly bracket (as in the first best) but also due to the distortion (the last term in the r.h.s.).

The optimal number of competing agent n^* is implicitly defined by the two inequalities in (13) where the left hand side is the cost increase of a unitary reduction of competing agents from n to $n-1$ and the right hand side is the cost reduction of a unitary increase of agents from n to $n+1$. The intermediate term is the effect on distortion in V_n induced by a unitary change (increase or reduction) of the number of agents in terms of quality. Increasing n^* by one more agent increases the distortion on quality with a negative effect on V_n that is larger than the efficiency gain on agents' cost. On the other hand, excluding one more firm from n^* generates an efficiency loss that has a more important effect on V_n than the reduction of distortion on quality. (13) shows that n^* is larger (smaller) the smaller (larger) is the distortion $c(q,x)/\delta^x$ so that $n^* \leq n_{FB}$.

The relation between optimal x^* and n^* is finally characterized by even stronger substitutability than with first best since the terms induced by the distortion in the previous first order conditions all point in the same direction of stronger substitutability.

Hence, from all these reasonings it follows that since from (11) $q^*(v)$ is increasing in v, then $n^*(v)$ and $x^*(v)$ are both decreasing in v.

(ii) Since a larger B increases V_N and helps satisfying (3) the principal increases as much as possible B up to $c(q,x)$, i.e. $B = c(q,x)$, so that and the program becomes

$$\begin{aligned} &\max_{q,x} \quad V_N \;\; \left(= s(q)\tfrac{1}{1-\delta} - \left\{E[\theta'(N)] + \theta_e(x) + k\right\}\tfrac{1}{1-\delta^x}\right) \\ &s.t. \\ &(3):\;\; i.e.\; (V_N - \underline{V}_N)\delta \geq c(q,x) \end{aligned}$$

Clearly, since V_N is the same objective function in the first best, if q_{FB}, x_{FB} satisfy (3) then the optimal C_N^* is fully efficient. This possibility will be illustrated below. If instead with q_{FB}, x_{FB} (3) binds then the optimal C_N^* is distorted with (3) necessarily binding. In this case, let $\lambda \geq 0$ be the Kuhn-Tucker multiplier of constraint (3). The optimality conditions for q and x are respectively

$$\frac{\partial V_N}{\partial q} = \frac{\partial c(q,x)}{\partial q}\frac{\lambda}{\delta(1+\lambda)}, \quad \frac{\partial V_N}{\partial x} = \frac{\partial c(q,x)}{\partial x}\frac{\lambda}{\delta(1+\lambda)}.$$

Since $\partial c(q,x)\backslash\partial q \geq 0$, $\partial c(q,x)\backslash\partial x \geq 0$ it follows that $q_N^* \leq q_{FB}$, $x_N^* \leq x_{FB}$.

The derivation of the optimal C_N^* would require to solve backward from $N = 1$ and increasing N. Although will not follow this approach here, it is worth noticing that $V_N \geq \underline{V}_N$ and, if at the optimum (3) does not bind for a given N, then it also does not bind at the optimum for $N' \leq N$ since $V_N - \underline{V}_N$ is decreasing in N.

Notice that, as long as (3) does not bind, then V_N^* is increasing in N. Consider first a N such that does not bind (3) so that $C_N^* = C_{FB}$. This implies that constraint (3) can be written as

$$\left\{E[\theta'(N-1)] + \theta_e(x_{N-1}^*) + k\right\}\frac{1}{1-\delta^{x_{N-1}^*}} - \left\{E[\theta'(N)] + \theta_e(x_N^*) + k\right\}\frac{1}{1-\delta^{x_N^*}} > c(q_N^*, x_N^*)/\delta \tag{14}$$

From Lemma 1 we can have two cases: either $x_{N-1}^* = x_N^*$ or $x_{N-1}^* > x_N^*$. First, let $x_{N-1}^* = x_N^*$ so that the previous expression reduces to

$$\Delta\theta(N-1)\frac{\delta}{1-\delta^{x_N^*}} > c(q_N^*, x_N^*).$$

Since the l.h.s. decreases in N, it exists an $\underline{N} \in \Im_+$ so that for $N \geq \underline{N}$ the previous inequality is violated and the constraint (3) must bind and for $N < \underline{N}$ the inequality is satisfied so that $C^*_N = C^{FB}$. Clearly, the larger is v, the higher is q^*_N and the smaller is $\underline{N}$. Second, consider now $x^*_{N-1} > x^*_N$ so that the first term in the l.h.s. of (14) is even larger than the second and the result holds *a fortiori.* Finally notice that since q^*_N is increasing in v, then $\underline{N}$ is a decreasing function in v.

The characterization of C^*_N for any $N \geq \underline{N}$ for which (3) binds goes as follows. When $N = \underline{N}$ the both $q^*_{\underline{N}}$ and $x^*_{\underline{N}}$ discretely jump respectively below q_{FB} and x_{FB} (i.e. the optimal quality and contract length with $N < \underline{N}$), as implied by the binding constraint (3). Furthermore, for $N > \underline{N}$, q^*_N and x^*_N must decrease in N. To see this consider first $N = \underline{N} + 1$ and let $\tilde{V}_{\underline{N}}$ be $V_{\underline{N}}$ where we replace the optimal $q^*_{\underline{N}}$ and $x^*_{\underline{N}}$ for $\underline{N}$ with q_N and x_N. We then have

$$V_N - \tilde{V}_{\underline{N}} = \Delta\theta(\underline{N})\frac{1}{1-\delta^{x_N}} \geq V_N - \underline{V}_{\underline{N}}$$

where the inequality comes from $\tilde{V}_{\underline{N}} \geq V_{\underline{N}}$ by definition. Using this with (3) we have, for $N = \underline{N}+1$,

$$\Delta\theta(\underline{N})\frac{1}{1-\delta^{x_N}} \geq V_N - \underline{V}_{\underline{N}} \geq c(q_N, x_N)/\delta$$

As argued above, since $\Delta\theta(.)$ is decreasing, q_N, x_N must be decreasing in N. This argument can be clearly replicated for any pair $N(> \underline{N})$ and $N'(= N+1)$ thus implying the result on q_N, x_N, with $\lim_{N\to\infty} x^*_N = 1$ and $\lim_{N\to\infty} q^*_N = 0$. ■

Proof of Corollary 1. Clearly for $N \leq \underline{N}(v)$ the relational contract C^*_N is the overall optimum since it corresponds to the first best. From now on we consider $N > \underline{N}(v)$. The buyer has three options: either open procurement getting V^*_N, or restricted procurement with $n > 1$ and a payoff V^*_n or she can set $n = 1$ (i.e. negotiation) and a payoff V^*_1.

For what stated in the proof of Proposition 1 it follows that if N is large then the distortion on open procurement can be very large and, in the limit, it obliges the principal to set $q = 0$ since

$$\Delta\theta(N)\frac{1}{1-\delta^{x_N}} \geq V_N - \underline{V}_{\underline{N}} \geq c(q_N, x_N)/\delta$$

and $\lim\limits_{N\to\infty} \Delta\theta(N) = 0$. Hence, since for large N the payoff V^*_n with restricted tendering does not depend on N, it follows that there exists a $\overline{N}(v)$ such that $V^*_n \geq V^*_N$ for $N \geq \overline{N}(v)$ with $\overline{N}(v)$ increasing in v since the larger is v the higher is $c(q,x)/\delta$.

Furthermore, when v is very small then principal's concerns for quality tend to disappear so that she prefer open tendering so that it must be $\overline{N}(v) = \underline{N}(v)$ when quality is not an issue, i.e. $v = 0$, and the principal wants maximal competition and $\overline{N}(v) - \underline{N}(v)$ is increasing in v.

If instead N is intermediate, i.e. $\in (\underline{N}(v), \overline{N}(v))$, the comparison is ambiguous because both procurement modes are characterized by distortions which cannot be compared in general. ■

Proof of Corollary 2. Suppose first the buyer cannot use B. Constraint (2) must sill bind at the optimum and the only difference between open and restricted procurement is the number of competitors: as in the proof of Proposition **??** the trade-off is that more competitors increase efficiency but also increase the distortion on quality.

Suppose now that the principal cannot use w. First notice that (1) is always satisfied. Hence, for any $n \leq N$ the principals's program consists in maximizing

$$V_n = [v(q) - \psi(q)] \frac{1}{1-\delta} - \left\{ E[\theta''(n)] + \theta_e(x) + k\, \mathcal{I}_{n>1} \right\} \frac{1}{1-\delta^x}$$

subject to (2), i.e.

$$\beta(n)\pi(n) \geq \frac{1-\delta^x}{\delta^x} [c(q,x) - B]$$

and (3), where the cost component now depends on the cost of the second most efficient agent, i.e. $E[\theta''(n)]$.

With open procurement, since a larger B has no effect on V_N and relaxes constraint (2), the principal optimally sets B so that (3) binds and (2) finally becomes

$$\beta(N)\pi(N) \geq \frac{1-\delta^x}{\delta^x} \left[c(q,x) - \delta \left(V_N - \underline{V}_N\right)\right].$$

In the case of restricted procurement and negotiation, the reasoning is similar except for $B = 0$ so that the principal maximizes V_n subject to (2), i.e.

$$\beta(n)\pi(n) \geq \frac{1-\delta^x}{\delta^x} c(q,x).$$

If (2) does not bind at the optimum, open procurement is preferred due to smaller costs, i.e. $E[\theta''(N)] < E[\theta''(n)]$ for any $n \in [1, N)$. However, if v is sufficiently high then (2) binds and implicitly defines the maximal implementable quality $q(x,n)$. Comparing the two constraints with N and n agents, two effects are at play. First with more agents the l.h.s. is smaller since both $\beta(.)$ and $\pi(.)$ are decreasing functions. Second, the positive effect of B with open procurement (that reduces the r.h.s. of constraint (2) by $(1-\delta^x)/\delta^{x-1}\ (V_N - \underline{V}_N))$ is decreasing in N. Hence, Proposition **??** holds true. Furthermore, the impossibility to use w reduces the l.h.s. (2) and makes the desirable effects on q of reducing the number of competing agents even stronger.

Finally, if the buyer can use neither B nor w, our previous analysis combines and the result holds. ■

Proof of Lemma 4. (i) Suppose first the principal cannot use w nor B and consider a relational contract so that the maximal implementable quality when agents do no collude is $q(n,x)$ implicitly defined by (7). A necessary condition for cooperation to be incentive compatible is $\pi(n) \geq \pi^*(n)$. Hence, from (7) the result immediately follows. Assume now that the principal can only use B but not w. Cooperation is stable if

$$[B + \beta(n)\pi(n)]\,\frac{\delta^x}{1-\delta^x} \geq D + \max\{B + \beta(n)\pi^*(n)\frac{\delta^{x^*}}{1-\delta^{x^*}}, c(q,x)\}$$

and the current supplier does not cheat on quality if

$$[B + \beta(N)\pi(N)]\,\frac{\delta^x}{1-\delta^x} \geq c(q,x) - B.$$

Comparing competition and cooperation in the case of open procurement, we notice that a unitary increase of B with cooperation relaxes the constraint on quality by $1 + \frac{\delta^x}{1-\delta^x}$ which is more that would happen with a similar increase of B with competition. Hence, ceteris paribus the principal can implement a larger quality with cooperation.

If the principal can use w but not B, cartel's stability constraint is

$$[w + \beta(n)\pi(n)]\,\frac{\delta^x}{1-\delta^x} \geq D + \max\{[w^* + \beta(n)\pi^*(n)]\,\frac{\delta^{x^*}}{1-\delta^{x^*}}, c(q,x)\}$$

which substituting $w^* = c(q^*,x^*)\frac{1-\delta^{x^*}}{\delta^{x^*}} - \beta(n)\pi^*(n)$ becomes

$$[w + \beta(n)\pi(n)]\,\frac{\delta^x}{1-\delta^x} \geq D + \max\{c(q^*,x^*), c(q,x)\}$$

Hence, since $D + \max\{c(q^*,x^*), c(q,x)\} \geq c(q,x)$, the incentive compatibility constraint for cooperation implies the supplier's incentive compatibility constraint for quality, i.e. $[w + \beta(n)\pi(n)]\,\frac{\delta^x}{1-\delta^x} \geq c(q,x)$ which is the same whether agents cooperate or not.

(ii) Assume that the principal can use both B and w. Incentive compatibility constraint of cooperation (8) then becomes

$$[w + B + \beta(N)\pi(N)]\,\frac{\delta^x}{1-\delta^x} \geq D + \max\{B, c(q,x)\}. \tag{15}$$

For our argument it suffices to consider here $\delta^x \geq 1/2$ so that a larger B actually strengthen cooperation. Constraints (15) and (9) can be put together thus leading to the single constraint,

$$[w + B + \beta(N)\pi(N)]\,\frac{\delta^x}{1-\delta^x} \geq \max\{D + \max\{B, c(q,x)\}, c(q,x) - B\}.$$

Suppose $c(q,x) - B \geq D + max\{B, c(q,x)\}$ which requires $c(q,x) \geq B$ because $D + max\{B, c(q,x)\} > 0$, so that the previous becomes $-B \geq D$: this is impossible since $D > 0$. Hence, it must be $c(q,x) - B \leq D + max\{B, c(q,x)\}$ and the relevant constraint is (15). This immediately shows that the constraint on quality is not a concern once cooperation is in place because it is itself implied by the cooperation's stability. Hence, comparing cooperation with competition a larger quality can be implemented in the former case. ■

Proof of Proposition 3. (i) Assume the principal cannot use w. First notice that (8) can be written as

$$[B + \beta(n)\tilde{\pi}(n)]\,\frac{\delta^x}{1-\delta^x} \geq D + \max\{B + \beta(n)\pi^*(n)\frac{\delta^{x^*}}{1-\delta^{x^*}} - [c(q,x) - B], B\} + c(q,x) - B$$

which implies the constraint on quality (9), i.e.

$$[B + \beta(n)\tilde{\pi}(n)]\,\frac{\delta^x}{1-\delta^x} \geq c(q,x) - B.$$

Hence, if v is large enough, the principal wants to implement a large q and then constraint (8) binds. As previously shown this, together with $\tilde{\pi}(n) \geq \pi^*(n)$ immediately imply that, for any B, the principal prefers inducing collusion so as to have a larger q. In addition, since with $n = N$ and colluding agents B is part of agents' profits, then for any $B > 0$ the previous result holds a fortiori.

Consider now the possibility that $n = 1$. In this case, the principal induces the largest implementable expected rent to the (single) supplier and then creates the strongest incentives for implementable quality q that allow to implement the largest quality q. Indeed, consider the rent of the single agent with $n = 1$ and let us compare it with the expected rent with $n > 1$ colluding agents. With probability $1/n$ the single agent with negotiation will have the same efficiency as the most efficient agent in the cartel of n agents, thus earning the same profit. With complementary probability it will not be as efficient but still earn a profit contrary to what would happen for that agent being in the cartel.

As a last point, imagine now that the principal also cares for efficiency so that she maximizes

$$V_n - \gamma E[\theta_{it} + \theta(x) + c(q,x)],$$

where $\gamma \geq 0$ is some weight to the expected cost of procurement. Consider then the case in which the value for quality is high so that the principal either procure with colluding agents or with negotiation. Since for given q and x the procurement cost with negotiation is

$$E[\theta_e + \theta(x) + c(q,x)] \geq E[\theta'(n) + \theta(x) + c(q,x)]$$

where the r.h.s. is the cost with colluding agents, then for sufficiently large γ the principal prefers procurement with colluding agents.

(ii) Finally, assume now the principal can use the fix fee w. Following the same steps as above, the relevant constraint with collusion is the cartel's stability so that

$$w = [D + \max\{B, c(q,x)\}] \frac{1-\delta^x}{\delta^x} - B - \beta(n)\tilde{\pi}(n)$$

(recall that with competition $c(q^*, x^*) = B^*$). This transfer is certainly larger than that with competing agents if the principal must set $B = 0$ (if B is not feasible or if $n < N$) since with competition for any $B \geq 0$ we have

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

Hence, if the principal cannot use B, then collusion is immediately dominated by competition.

If $B > 0$ then the comparison of the two previous transfers shows a positive effect of B with collusion that is absent with competition: with colluding agents the principal actually pays B which allows to reduce w to all N suppliers with a net benefit in V_N for the principal equal to $\frac{1}{1-\delta^x} B(N-1)$:

$$\begin{aligned} V_N &= [v(q) - \psi(q)] \frac{1}{1-\delta} - \left\{E[\theta'(N)] + \theta(x) + k\right\} \frac{1}{1-\delta^x} + \\ &\quad -N \frac{1}{1-\delta^x} D - \max\{B, c(q,x)\} \frac{N}{\delta^x} + \frac{1}{1-\delta^x} B(N-1). \end{aligned}$$

Hence, the net effect of collusion is potentially ambiguous. If the derivative of V_N w.r.t. B is negative then $B = 0$ and again collusion is dominated, as previously shown. Suppose $B \leq c(q,x)$ so that the derivative of V_N is positive. The benefit of using $B > 0$, i.e. $\frac{1}{1-\delta^x} B(N-1)$, has to be confronted with the cost of collusion, i.e. the increase of w by $D\frac{1-\delta^x}{\delta^x}$ with a cost in V_N equal to $-N\frac{1}{1-\delta^x} D$. If N is sufficiently large, then following the same steps as in the proof of Proposition 1 part (ii), B is small and then $\frac{1}{1-\delta^x} B(N-1) \leq N\frac{1}{1-\delta^x} D$: the cost prevails. If instead N is small, then inducing collusion may be optimal if $(N-1)/N > D/B$. ■