Relational incentive contracts with collusion

Marta Troya-Martinez and Liam Wren-Lewis*

January 30, 2017

Abstract

This article explores how the possibility of collusion affects relational contracts. Responsibility for a contract is delegated to a supervisor who cares about both production and kickbacks paid by the agents, neither of which are explicitly contractible. We characterize the optimal supervisor-agent relational contract and show that the relationship between joint surplus and production is non-monotonic. Delegation may benefit the principal when relational contracting is difficult by easing the time inconsistency problem of paying incentive payments. For the principal, the optimal supervisor has incentives that are partially, but not completely, aligned with her own.

JEL classifications: D73, D86, L14.
Keywords: Relational contracts, delegation, corruption.

*Troya-Martinez: New Economic School, 100A Novaya Street, Moscow 143026, Russia (email: mtroya@nes.ru); Wren-Lewis: Paris School of Economics, 48 Boulevard Jourdan, 75014 Paris, France (email: liam.wren-lewis@parisschoolofeconomics.eu). We are grateful to Heski Bar-Isaac, Simon Board, Ernesto Dal Bó, Aaron Bodoh-Creed, Mikhail Drugov, Matthew Ellman, Florian Englmaier, Matthias Fahn, Frederico Finan, William Fuchs, Ben Hermalin, Elisabetta Iossa, Jonathan Levin, Michael Katz, Rocco Macchiavello, Jim Malcomson, David Martimort, Gerard Padró i Miquel, Orie Shelef, Giancarlo Spagnolo, John Vickers and the participants of AEA, ASSET, BGSE Summer Forum, CSAE, CMPO, EEA, ES World Congress, GREThA, and ISNIE conferences, the Workshop on Relational Contracts, and the seminars at HSE, LMU Munich, NES, Oxford, Paris I, PSE, SSE, UAB and UC Berkeley for many insightful comments. Any remaining errors are our own.

A wide range of important economic activities depend on self-enforcing contracts.[1] Responsibility for these contracts is frequently delegated to intermediaries; firms delegate to managers, governments to bureaucrats. Recent papers suggest that, although delegation can improve productivity, it is often held back by a lack of trust (Bloom, Sadun and Van Reenen, 2012; Bloom et al., 2013). In particular, organizations fear that intermediaries may extract kickbacks in the form of bribes or nonmonetary private benefits. These forms of collusion cannot be legally enforced, and they are therefore also sustained by relational contracts.

If intermediaries can collude with agents, how and when should relational contracts be delegated? This paper answers this question by extending a standard principal-agent relational contracting model to include an intermediary supervisor. We characterize the optimal bilateral relational contract between the supervisor and each agent and show that, if there is a cap on the payments that the supervisor can authorize, effort can be a nonmonotonic function of discounted surplus. This differs from standard models of relational contracting, where greater surplus within the relationship leads to greater agent effort (Malcomson, 2013). When the supervisor is unconstrained by the payment cap, greater surplus sustains greater payments to reward the agents, inducing higher kickbacks and also greater effort. But, when the supervisor is constrained by the payment cap, she may pay the agents even when they are unsuccessful, increasing kickbacks but reducing effort. Thus, depending on the surplus in each supervisor-agent relationship, there may or may not be a trade-off between encouraging production and reducing collusion.

A further important contribution of the paper is to analyze the costs and benefits of delegation for the principal. We find that, because the supervisor cares less than the principal about payments to the agents, delegation enhances credibility. A supervisor who cares too little about payments, however, will overpay in exchange for kickbacks. The principal therefore faces a trade-off when deciding how much the supervisor's payoffs should

[1]For instance, Antràs and Foley (2015) and Macchiavello and Morjaria (2015) provide evidence that relational contracts play an important role in international supply chains, while Board (2011) and Spagnolo (2012) attest to their value in public procurement. Gibbons and Henderson (2013) and Blader et al. (2015) argue that variation in effective relational contracts within firms can explain significant differences in performance.

be aligned with her own. Overall, the possibility of collusion makes delegation costly for the principal, but when relational contracting is difficult this cost may be more than compensated for by the supervisor's greater credibility.

The paper begins in Section 1 by motivating our model through documenting the wide relevance of relational contracts that sustain both production and collusive side payments. We give examples of the kind of trade-offs that exist and highlight suggestive evidence that kickbacks can help to sustain productive relational contracts.

Section 2 then sets out the model, which is an extension of a moral hazard version of Levin (2003). In the model, a supervisor and a number of agents have bilateral relational contracts over an infinite number of periods. A principal sets the parameters of this interaction at the beginning of the game and then takes no further action. Every period, each agent exerts continuous hidden effort towards producing a binary output. The expected aggregate output is then split between the supervisor and principal.[2] In return, each agent receives compensation that is partly at the discretion of the supervisor. The part that the supervisor has control over - the 'bonus' - is subject to a cap. The supervisor and agents can also exchange noncontractible side payments.[3] To our knowledge, this is the first paper to build a model where productive and collusive relational contracts co-exist.

A first insight from the model is that the two parts of the relational contract interact with each other in important ways. There is a positive interaction, as the expected stream of future kickbacks allows the supervisor to credibly promise higher bonuses, and these bonuses may then be used to motivate greater effort. But there is also a negative interaction, as the supervisor must trade off inducing effort and sustaining bribery when self-enforcement is a binding constraint.

In Section 3 we then characterize optimal relational contracts between the supervisor and the agents. Stationary contracts are optimal and the supervisor may motivate effort using variation both in bribes and in bonuses.

[2]An important assumption in our model is that the supervisor can receive a share of the profit directly, but the agents cannot. We discuss potential justifications of this assumption at the end of Section 2. In Section 5.2.3 we show for which parameter range it is optimal not undertake profit-sharing with the agents.

[3]The terms side payments, bribes, and kickbacks are used interchangeably.

The choice of motivation tool and the effort that results depend on the discounted surplus that is shared between the supervisor and each agent. When surplus is very high, self-enforcement is not a problem, and the supervisor will always authorize the maximum possible bonus and motivate any effort through variation in bribes. When surplus falls below a certain level, the optimal contract may also involve variation in bonuses. This is because the supervisor only pays for part of the cost of any bonus, and it is therefore more credible for the supervisor to motivate effort through bonuses rather than bribes. Indeed, when surplus is low, bribes will not be used to motivate effort.

A notable result that follows is that the agents' effort is nonmonotonic in the joint supervisor-agent surplus. This is because increasing surplus can have two potentially conflicting effects on effort: it raises the amount of effort it is possible to induce, but it may also change the supervisor and agents' marginal benefit of effort. When surplus is low, bonuses are used to motivate effort, and hence one benefit of higher effort is that it increases the expected bonus. When surplus is high, bonuses are paid regardless of output, and hence this benefit of effort disappears. Thus the supervisor and agents desire a higher level of output when surplus is lower. Overall, an increase in surplus can lead to a decrease in effort.

This result has important implications for policies designed to reduce collusion in contexts where relational contracts are valuable. Many such policies involve disrupting relational contracts, for instance by encouraging competition or increasing personnel rotation. The results of our theoretical analysis suggest that, in some circumstances, weakening supervisor-agent relations may simultaneously cut corruption and improve output. In other circumstances, however, there will be a trade-off.

Section 4 then analyzes how and when the principal should delegate. We allow the principal to set the fixed transfer paid to the agents, the limit on discretionary bonuses, and the extent to which the supervisor's preferences are aligned with her own. We find that the principal should choose the supervisor's preferences to be partly, but not completely, aligned with her own. More aligned preferences make supervisor-agent relational contracting more difficult, and hence the principal has to give them more surplus. Indeed, a supervisor with exactly aligned preferences will never be

optimal, because reducing effort slightly below first best is a second-order cost for the principal, but giving up surplus is first-order. On the other hand, less aligned preferences make the supervisor more tempted to pay bonuses regardless of output. At the limit, a supervisor who doesn't care at all about profits will induce no effort. The optimal supervisor therefore always lies somewhere between the two extremes.

We then show that delegating to a supervisor who may collude can sometimes, but not always, improve the principal's payoff. The supervisor has a comparative advantage in enforcing relational contracts because she has more credibility when paying promised bonuses. She cares less about making payments and yet values the relationship with the agents because of the expected stream of future kickbacks. Looking at it another way, if the supervisor reneges on her promises, then the agents can punish her by withholding the kickback. If principal-agent relational contracting is difficult, then delegating induces higher effort and increases the principal's payoff. On the other hand, if relational contracting is easy, then the principal prefers not to delegate in order to avoid having to share part of the surplus with the supervisor.

We consider a number of alternative specifications and extensions in Section 5 of the paper. These include giving the supervisor a more general payoff function and making side payments costly. We show that giving the principal extra instruments can allow kickbacks to be replaced by direct payments and makes delegation more advantageous, but that first-best effort will only be induced if we allow seemingly unrealistic contracts. Making side payments more costly generally benefits the principal, leading to greater delegation, less profit sharing and greater production. Finally, Section 6 concludes by considering avenues for future theoretical and empirical work. Mathematical proofs of all lemmas and propositions are then given in the appendix.

This article bridges two large theoretical literatures - that on relational contracts and that on collusion in hierarchies. Models in the relational contracting literature have typically not considered hierarchies (see Malcomson, 2013, for a survey). A recent exception is the paper by Fong and Li (2016). In their model, the supervisor carries out subjective performance evaluations of the agent. They show that by garbling the evaluations

intertemporally, the supervisor can relax the self-enforceability problem thereby benefiting the principal. The possibility of collusion is however not the focus of their analysis and hence they conclude that *"formal study of how collusion affects relational contracts is an interesting line of future research"*.

Collusion between supervisors and agents has been the focus of a large literature including seminal works by Tirole (1986), Milgrom (1988) and Kofman and Lawarree (1993). A few papers have studied how the need for self-enforcement impacts such collusion - see, for instance, Martimort (1999), Martimort and Verdier (2004) and Dufwenberg and Spagnolo (2015) - but these do not consider any interaction with other commitment problems.[4]

Our model also relates to a literature investigating how delegation to an intermediary can solve commitment problems.[5] Within this literature, the paper closest to ours is that of Strausz (1997). In a static framework, he considers how delegating monitoring to an intermediary may benefit the principal through a 'commitment effect' and then studies the impact of collusion. The model differs from ours in a number of ways. For instance, he assumes that collusion is automatically enforced and that the supervisor can only trigger bonuses when the agent performs the (binary) action well. Given these different assumptions, he finds that potential supervisor-agent collusion has no impact on the principal's payoff, contrary to our results.

A recent paper that considers how delegated cooperation can be maintained is that of Hermalin (2015). He builds a model whereby 'wining and dining' helps sustain a productive relationship between two firms' managers, and he also considers managers colluding against their principals.

[4]Olsen and Torsvik (1998) find that potential supervisor-agent collusion can mitigate a commitment problem, but their model differs from ours by studying adverse selection rather than moral hazard. Also relatedly, Calzolari and Spagnolo (2009) look at the complementary between relational contracts and tacit collusion between firms, where side transfers are not possible.

[5]Earlier papers in this literature include Vickers (1985) and Katz (1991). Spagnolo (2005) uses a repeated game framework to show how delegation can enhance the enforcement of tacit collusion between firms. Like much of this literature, we assume that the intermediary's payoff function is observed by the agent. Our results may, however, be less vulnerable to principal-supervisor renegotiation than those of other papers, since in our context part of the supervisor's payoff comes directly from the agent - see Section 5.2.1 for more details.

A key difference between this paper and ours is that collusion is not sustained through relational contracting and only occurs when side payments are costless. As a result, allowing cross-firm managerial rewards always benefits the principals.

1 Examples of 'dual' relational contracts

We begin by motivating our investigation with examples of two types of relationship that frequently sustain both productive and collusive implicit contracting: relationships between firms delegated to purchasing managers and firm-employee relationships delegated to supervisors. We document evidence on how the two parts of these relational contracts may interact and demonstrate the existence of trade-offs that result.

1.1 Procurement and relationships between firms

It is now well established that relational contracts play a key role in transactions between organizations, including inter-firm trade and public procurement. A typical example is the purchase of goods where it is difficult to observe quality before buying. In this case, the purchaser may rely on a relational contract, inducing the seller to produce high quality by threatening partial nonpayment or the termination of the relationship.

Many procurement relationships are delegated to intermediaries, and it is well known that such delegation carries risks of kickbacks or other collusive behavior. A typical example is a purchasing manager who has discretion regarding prices paid to suppliers. A study of Indian firm-owners by Bloom et al. (2013, p. 40) notes that many *"did not trust non-family members. For example, they were concerned if they let their plant managers procure yarn they may do so at inflated rates from friends and receive kickbacks"*.[6] Indeed, Lambsdorff and Teksoz (2005, p. 139) argue that delegating responsibility for relational contracts is particularly vulnerable to corruption because *"pre-existing legal relationships can lower transaction costs and serve as a basis for the enforcement of corrupt arrangements"*.

[6]Similarly, Bloom, Sadun and Van Reenen (2012, p. 1667) find multinationals decentralize less in low-trust environments, and argue that CEOs *"worry about the plant manager taking bribes from equipment sellers"*.

Cole and Tran (2011) give evidence that these two aspects of relational contracts are interlinked. They describe kickbacks made by two firms to intermediaries within organizations that they supply. In one case, they note that relational contracts are needed because quality is not contractible and hence *"the supplier allows the client to hold back roughly 20 percent of the contract value until ... the client is satisfied that the product meets the specified quality. The kickback is paid only after all contract payments have been made."* (p. 411). In another case, the agent *"usually specifies the kickback amount in advance but typically does not start paying until the first deposit is made"* (p. 420). In both cases the ordering of payments suggests that kickbacks are partly being used as an enforcement device to enhance the intermediary's credibility.

1.2 Labor relations and organizational structure

A large portion of the relational contracting literature has focused on labor relations within organizations. Employees are frequently rewarded for effort with promotions, wage increases, or bonuses based on unverifiable subjective performance evaluations rather than contracted measures of output.

In many organizations these relational contracts are delegated to intermediary managers who have substantial control over incentives. The risk of collusion between intermediaries and employees in this setting is well known; Milgrom (1988), Fairburn and Malcomson (2001), and Thiele (2013), for instance, each consider the possibility of employees engaging in wasteful collusion or 'influence activities'.[7] As an example, Nkamleu and Kamgnia (2014) document that in African governments per-diems are *"mainly given to provide financial incentives to employees in order to increase their motivation"* (p. 4) but managers may *"expect the staff member to share or kickback a portion of the per diem"* (p. 12). Indeed, Rasul and Rogger (2016) find evidence in Nigeria that influence activities lead subjective performance evaluations to damage production.

[7]These papers implicitly assume that such manager-employee collusion can be automatically enforced, and hence do not explore how this behavior relates to relational contracts. Thiele (2013) considers a principal who may operate a relational contract with the agent, but assumes that delegation to a corruptible supervisor results in all contracts becoming court-enforceable.

The dual nature of relational contracts within organizations leads to conflicting implications regarding their value. For instance, Francois and Roberts (2003) argue that factors enabling relational contracts within firms increase employee productivity and innovation. On the other hand, Martimort and Verdier (2004) argue that the same factors increase the ability of supervisors and employees to collude and hence dampen economic growth. Consistent with this view, Khanna, Kim and Lu (2015) find that greater connections between CEOs and top executives increase the risk of corporate fraud. Clearly both views are possible, but it is difficult to evaluate the potential trade-off without understanding how the two parts of the relational contract interact.

2 The model

Our model contains a principal, a supervisor, and N identical agents. We refer to the principal and supervisor as female and the agents as male. We introduce the following changes to the standard principal-agent model of Levin (2003). The principal sets three key parameters at the beginning of the game, delegates responsibility for managing the contracts to the supervisor and then takes no further action. Contracts between the supervisor and agents are bilateral, so we can treat the relationship between the supervisor and each agent as a separate game (Levin, 2002).

The three parameters set by the principal are the wage w that each agent receives, a cap $\overline{b}$ on the size of the bonus that the supervisor can disperse to each agent, and a proportion α of profit that is given to the supervisor. These parameters apply for each period and cannot vary over time or as a function of output.[8] After these parameters are set, the supervisor and agents interact repeatedly over an infinite horizon of discrete periods.

The timeline for each period for a given agent is shown in Figure 1. The supervisor first proposes a compensation package and a set of side payments to each agent. Each agent either accepts or rejects - let $d_{it} \in \{0, 1\}$ denote

[8]Section 5.2.4 considers the possibility of allowing the wage to vary over time. Another way in which the principal could improve her payoff would be to replace the bonus cap with some sort of cap on 'average' bonuses. In practice, however, such a restriction may be too complex and subject to renegotiation, since the principal would like to 'reset' the bonus cap after several periods of high output.

agent i's decision. If an agent accepts, then he chooses an effort $e_{it} \in [0,1]$ incurring a cost $c(e_{it})$, where $c(0) = 0$, $c'(0) = 0$ and $c''(\cdot) > 0$. The agent's effort generates a binary stochastic output $y_{it} \in \{0, y\}$ where $0 < y$. The output is high ($y_{it} = y$) with probability e_{it} and low ($y_{it} = 0$) otherwise. If an agent rejects, then no output is produced and the agent recieves a per-period payoff of $\underline{u}$. Collective output Y_t is then the sum of these individual outputs and some noise, i.e. $Y_t = \sum_i y_{it} + \epsilon_t$, where ϵ_t iid with $\mathbb{E}[\epsilon] = 0$.

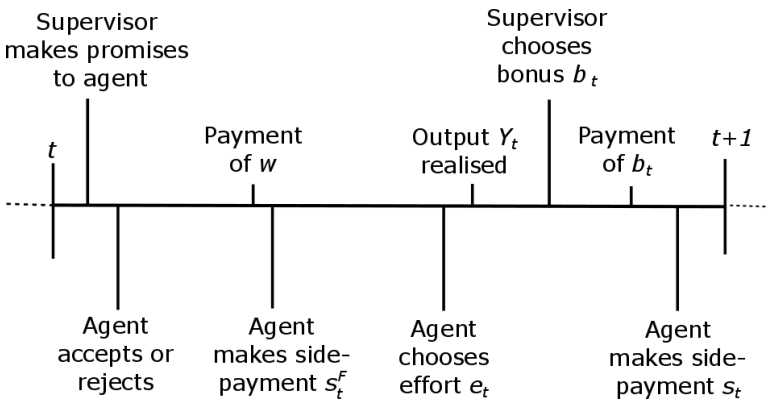


Figure 1: Timeline of period t in supervisor-agent game

Each agent's compensation package consists of the fixed payment w and a discretionary payment $b_{it} \in [0, \overline{b}]$ that can depend on output y_{it}; let $W_{it} = w + b_{it}$ denote the total payment made. In the context of procurement relationships, we can think of w as the upfront payment and b_{it} as the payment made after inspection of quality y_{it}.[9]

In addition to the compensation package, the supervisor also suggests a package of side payments that will be made from the agent to the supervisor. The kickback is paid in two parts: the first part, s_{it}^F, is paid before output is realized, while the second part, s_{it}, is paid after output is realized. Let $S_{it} = s_{it}^F + s_{it}$ denote the total side payment made.

The total profit $Y_t - \sum_i W_{it}$ is split between the principal and supervisor, who receive shares $1 - \alpha$ and α respectively, where $\alpha \in [0,1]$.[10] All players

[9]We assume that bonus payments b_{it} can be anywhere within the range $[0, \overline{b}]$, but it would not change our results if the supervisor could only choose between $b_{it} = 0$ and $b_t = \overline{b}$. In this case, mixed strategies would allow her to promise intermediate values in expectation.

[10]We hence assume that the principal and supervisor receive an equivalent to zero

have the same discount factor $\delta \in (0,1)$. The expected payoff functions can thus be written as follows:

$$\Pi_t = \sum_{i=1}^{i=N} \pi_{it} = \sum_{i=1}^{i=N} (1-\delta)\mathbb{E}\left[\sum_{\tau=t}^{\infty} \delta^{\tau-t}(1-\alpha)(y_{i\tau} - W_{i\tau})\right]$$

$$V_t = \sum_{i=1}^{i=N} v_{it} = \sum_{i=1}^{i=N} (1-\delta)\mathbb{E}\left[\sum_{t=\tau}^{\infty} \delta^{\tau-t}[\alpha(y_{i\tau} - W_{i\tau}) + S_{i\tau}]\right]$$

$$u_{it} = \mathbb{E}\left[(1-\delta)\sum_{\tau=t}^{\infty} \delta^{\tau-t}\left\{d_{i\tau}[W_{i\tau} - S_{i\tau} - c(e_{i\tau})] + (1-d_{i\tau})\,\underline{u}\right\}\right]$$

where π_{it} and v_{it} are the expected payoffs that the principal and supervisor derive in their relationship with agent i.

The assumption that the supervisor receives a fraction α of the total profit $Y_t - \sum_i W_{it}$ is important.[11] In most of the paper, we assume that agents are not similarly motivated. Such a disparity may result from the supervisor being intrinsically motivated or entitled to more information than the agents. A private company, for instance, may not want profit to be revealed to low-level employees or actors outside of the company. Alternatively, we could consider that Y_t is verifiable, but that sharing profits with agents is ineffective due to there being a large number of agents. Indeed, in Section 5.2.3 we allow for formal contracting on aggregate output with the agents and show that, for sufficiently large N, the supervisor and principal will never wish to use such contracts.

The information structure is one of moral hazard. Effort is the agent's private information, while the individual output and agent's compensation are observed by both the supervisor and the particular agent. Agents cannot observe the individual output of the other agents, and the principal only observes collective output.

effort as their outside option, as in Levin (2002). This ensures our results are not driven by different relative valuations of the outside option by the principal and supervisor.

[11] We choose this form of supervisor payoff as it is the simplest that generates the key insights of the paper. Section 5.2 discusses alternatives, including where the supervisor receives a wage directly and has different weights on W_t and Y_t, and shows that the main results still hold.

3 Optimal supervisor-agent contracts

In this section, we are temporarily setting aside the actions of the principal and consider how the supervisor and each agent interact for any given parameters w, $\overline{b}$ and α. This is useful as a first step in understanding the principal's optimal behavior and also provides insights for contexts where principals may not be optimizing over these variables. Optimal strategies for the principal are then considered in Section 4.

We consider bilateral relational contracting between the supervisor and each agent. Since the relationships between the supervisor and each agent are technoligically independent of one another, we can thus treat the supervisor's relationship with each agent as a seperate game Levin (2002). We therefore omit the i subscripts from the notation for clearer exposition.

We begin this section by considering the benchmark case when $\alpha = 1$, since this is equivalent to a standard principal-agent model. We then consider the more general case when $\alpha < 1$ and solve for the optimal supervisor-agent contracts.

3.1 Optimal supervisor-agent contracts when $\alpha = 1$

If $\alpha = 1$, then our model is equivalent to one of principal-agent relational contracting such as Levin (2003). Side payments, bonuses, and wages are substitutable tools for making transfers between the the supervisor and each agent. The cap on bonuses and the invariance of wages have no consequence because side payments are a perfect substitute. Moreover, the supervisor receives all of the profit and hence has the same payoff function as the principal would without delegation. We can thus treat the results of this case as the 'no delegation' benchmark and refer to them as the outcome of direct principal-agent relational contracting.

If the supervisor and agent could contract explicitly on y_t, it would be optimal to induce the value of effort e_t that maximizes the joint surplus, $ye_t - c(e_t)$. Defining this first-best effort as e^{FB}, we then have $c'(e^{FB}) = y$.

When the supervisor and agent cannot explicitly contract on y_t, a self-enforcing contract is needed. We follow the definition of a self-enforcing contract given by Levin (2003) and similarly define a self-enforcing contract as optimal if no other self-enforcing contract generates higher expected

surplus for the supervisor and agent. Levin (2003) shows that, if we are concerned with optimal contracts, then there is no loss of generality in focusing on stationary optimal contracts. Moreover, any optimal contract will have effort constrained by the following inequality:

$$c'(e) \leq \frac{\delta}{1-\delta}(ey - c(e) - \underline{u}) \tag{1}$$

Note that the right hand side of the this inequality is the discounted joint surplus generated by the supervisor-agent relationship. When the future relationship is not valuable enough, the supervisor cannot credibly pay a bonus large enough to implement the first-best effort. Instead, the effective reward for high output will be the largest that can be credibly promised. Effort will therefore be increasing in the discounted joint surplus.

3.2 Optimal supervisor-agent contracts when $\alpha < 1$

In this section, we first show that we can still restrict our attention to stationary contracts that maximize joint supervisor-agent surplus. We then outline the key constraints that will potentially bind in any optimal supervisor-agent contract. This allows us to derive the main proposition in this section, which characterizes the optimal contract as a function of the discounted joint surplus. Finally, we examine the varying ways in which effort is motivated and detail how the relationship between effort and surplus is nonmonotonic.

A first point to note is that, when $\alpha < 1$, the surplus generated directly depends on the compensation scheme. This is because the supervisor only pays for part of the payment W_t that each agent receives. If explicit contracting on y_t were possible, the supervisor and agents would maximize their joint surplus by setting bonuses at the bonus cap $\overline{b}$ regardless of output and then use side payments to induce an effort level e_{SA}^{FB}, where $c'(e_{SA}^{FB}) = \alpha y$.

Side payments can be used to divide surplus between the supervisor and each agent. We can therefore focus on relational contracts that generate the largest possible supervisor-agent surplus. We follow Levin (2003) in defining a self-enforcing contract as strongly optimal if the continuation

contract is optimal for all potential histories, even those off-equilibrium.[12] We then obtain the following lemma:

Lemma 1. *If an optimal contract exists, there are stationary contracts that are strongly optimal.*

The intuition behind this stationarity result is that any variation in promised continuation values can be transferred into side payments in the same way that, in the principal-agent case studied by Levin (2003), any variation can be transferred to bonus payments.

We therefore focus on stationary contracts and drop the t subscripts on our variables. Let b_h be the bonus when output is high, b_l that when output is low, and define s_h and s_l similarly. Then effort will be determined by the following binding incentive compatibility constraint:

$$c'(e) = b_h - b_l - s_h + s_l \qquad (IC)$$

We define $g(e, b_h, b_l)$ as the expected supervisor-agent surplus in any stationary contract that has bonuses b_h and b_l and induces effort e. This is given by the following equation:

$$g(e, b_h, b_l) = \alpha e y + (1 - \alpha)(w + e b_h + (1 - e) b_l) - c(e) - \underline{u}$$

Note that, within stationary contracts, there are two ways that the supervisor can motivate effort: through variation in bonuses or in bribes. The following lemma shows that bonuses will never be negative and, if bribes or bonuses vary as a function of output, then they will do so in a way that encourages effort.

Lemma 2. *In any optimal contract, bonuses are always nonnegative, i.e.* $b_h \geq 0$ *and* $b_l \geq 0$*. Moreover, bonuses are weakly higher when output is high* ($b_h \geq b_l$) *and side payments are weakly lower* ($s_h \leq s_l$)*.*

[12] The concept of strong optimality defined by Levin (2003) is an equilibrium selection device that implicitly assumes that renegotiation can only take place if there are potential Pareto improvements; see Goldlücke and Kranz (2013) for how this condition relates to other renegotiation-proof concepts. Miller and Watson (2013) construct an alternative condition that assumes players bargain within each period and show that this typically involves suboptimal play after deviation.

If the supervisor wants to take surplus from the agents, then she prefers to do so using bribes rather than bonuses. This is because bribes and bonuses are equivalent for the agents, but the supervisor captures the whole value of any bribes given.

The need for the contract to be self-enforcing can be expressed in terms of dynamic enforcement constraints which demand that the future benefits of continuing the relationship are larger than the static gains from reneging on promises. Lemma 2 pins down the binding dynamic enforcement constraints that potentially bind. Since bonuses are never negative, only the supervisor has a reason to deviate when it comes to the bonus payment. This temptation will be greatest when output is high, as this is when the bonus is greatest. On the other hand, only the agents may wish to deviate from paying the agreed side payments, because if the supervisor does not wish to pay the side payment, she already would have deviated by not paying the bonus. The agents will be most tempted to renege when output is low, as this is when the side payment is greatest. We therefore need to concentrate on the two following dynamic enforcement constraints:

$$(1-\delta)\left(-\alpha b_h + s_h\right) + \delta v \geq 0 \qquad (DE_S)$$

$$-(1-\delta)s_l + \delta u \geq \delta \underline{u} \qquad (DE_A)$$

From these constraints, we can see that variation in bonuses is easier to sustain than variation in bribes. Increasing b_h by 1 only tightens (DE_S) by an amount α, but increasing s_l or decreasing s_h by 1 tightens the constrains by 1. In other words, motivating effort through bonuses is easier than motivating effort through bribes.

Summing (DE_S) and (DE_A) together and substituting in (IC) then gives us the following constraint:

$$c'(e) + \alpha b_h - [b_h - b_l] \leq \frac{\delta g(e, b_h, b_l)}{1-\delta} \qquad (IC - DE)$$

Comparing this to (1), the equivalent in the principal-agent case, we see that the requirement for contracts to be self-enforcing has a more complex impact in the supervisor-agent game. In particular, as the surplus in the relationship decreases, a reduction in effort is now only one possible effect.

The supervisor and agent may instead choose to keep effort constant and increase the difference in the bonuses. This makes relational contracting easier since it is more credible for the supervisor to induce effort using bonuses rather than bribes.

Define δ^{FB} as the critical level of δ at which the supervisor and agents can implement their first-best contract, i.e. $\frac{\delta^{FB}}{1-\delta^{FB}} = \frac{\alpha y + \alpha \bar{b}}{g(e_{SA}^{FB}, \bar{b}, \bar{b})}$. Then we obtain the following lemma:

Lemma 3. *If $\delta \leq \delta_{FB}$, then $(IC - DE)$ is binding.*

The ability to transfer utility through side payments ensures that there cannot be a second-best optimal contract where one of the dynamic enforcement constraints has slack. Hence, in any optimal contract that does not achieve first best, both (DE_S) and (DE_A) will be binding, and therefore so will $(IC - DE)$.

From $(IC - DE)$, we can see that changes in discounted surplus may affect effort, the bonus b_h, and the extent to which effort is induced through variation in bonuses rather than bribes. The following proposition characterizes the optimal contract as a function of δ and shows that the relationship between effort and δ is nonmonotonic. We can think of δ as a determinant of the potential discounted joint surplus, and indeed the proposition could be written similarly in terms of the wage w or the agents' outside option $\underline{u}$.

Proposition 1. *Agents' effort may be a nonmonotonic function of the surplus. In particular, there exist values δ^L and δ^H such that the optimal supervisor-agent relational contract can be characterized as follows:*

- ***High surplus:*** *If $\delta \geq \delta^H$, then bonuses are not used to induce effort, i.e. $b_h = b_l$, and effort is weakly increasing in δ.*
- ***Intermediate surplus:*** *If $\delta^H > \delta > \delta^L$, then both bribes and bonuses are used to induce effort, i.e. $b_l < b_h$ and $s_h < s_l$, and bonuses are at the cap when output is high, i.e. $b_h = \bar{b}$. If $b_l > 0$, then effort is decreasing in δ, and otherwise it is increasing in δ.*
- ***Low surplus:*** *If $\delta \leq \delta^L$ then bribes are not used to induce effort, i.e. $s_h = s_l$, and effort is weakly increasing in δ.*

The basic intuition behind the nonmonotonicity result can be understood as follows. When surplus is high, the relationship can sustain both large unvarying bonuses and a large variation in bribes to induce effort. When surplus falls below a certain level, the supervisor replaces some of the variation in bribes with variation in bonuses, since these are easier to sustain. Doing so means that effort benefits the supervisor and agents more, since high output now triggers higher bonuses. Lower surplus makes inducing effort more difficult, but this is more than compensated for by the increase in the value of effort for the two parties.

In order to better understand the nature of the optimal supervisor-agent contract, we now briefly detail the three cases outlined in Proposition 1. We also depict in Figure 2 the optimal contract for particular parameter values when $c(e) = \frac{1}{2}ce^2$ and the supervisor and agents have equal bargaining powers, i.e. $u = v = g/2$. Figures 2a and 2b plot the bonuses and kickbacks as a function of δ, with the latter including the expected bribe $s^F + es_h + (1-e)s_l$. Figure 2c then plots the induced effort levels, and for comparison we also include the effort level that would be exerted without delegation. Finally, Figure 2d plots the players' payoffs and the supervisor-agent surplus g.

3.2.1 High surplus

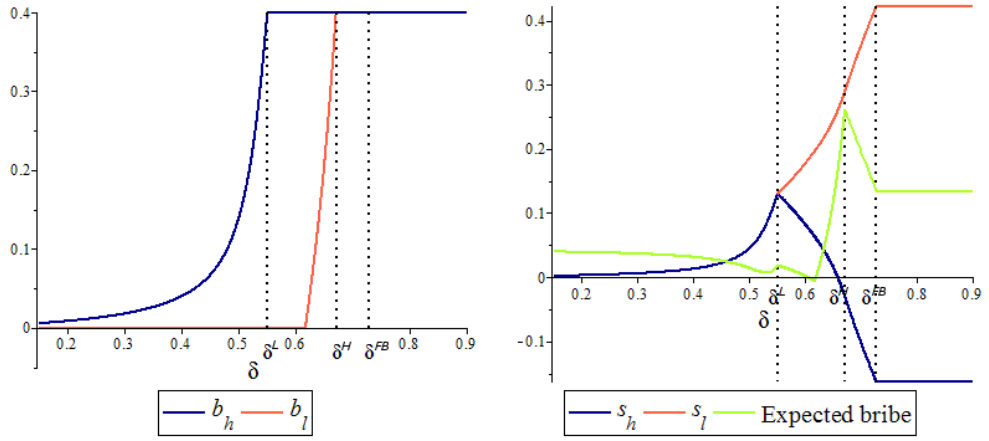
When surplus is slightly below the level that allows the supervisor and agent's first-best contract, effort will be below first best but bonuses will remain at the cap regardless of output. In particular, since $(IC - DE)$ is binding, effort will be determined by the equation $c'(e) = \frac{\delta g(e,\overline{b},\overline{b})}{1-\delta} - \alpha\overline{b}$. Effort is reduced before bonuses because, when $e = e^{FB}_{SA}$, a marginal reduction in effort leads to a second-order reduction in supervisor-agent surplus, while the cost of reducing the bonuses is first-order. There will thus always be a range of surplus for which the optimal contract involves $b_l = b_h = \overline{b}$ and $e < e^{FB}_{SA}$. Hence $\delta^H < \delta^{FB}$.

When surplus falls further, what happens depends on the relative value that the supervisor places on output, α. If α is low, then she will continue to cut effort rather than bonuses until no effort is sustainable. In this case $\delta^H \leq \delta^L$ and there is no 'intermediate' range of surplus. If α is high, then

Figure 2: Optimal supervisor-agent contract as a function of the discount factor

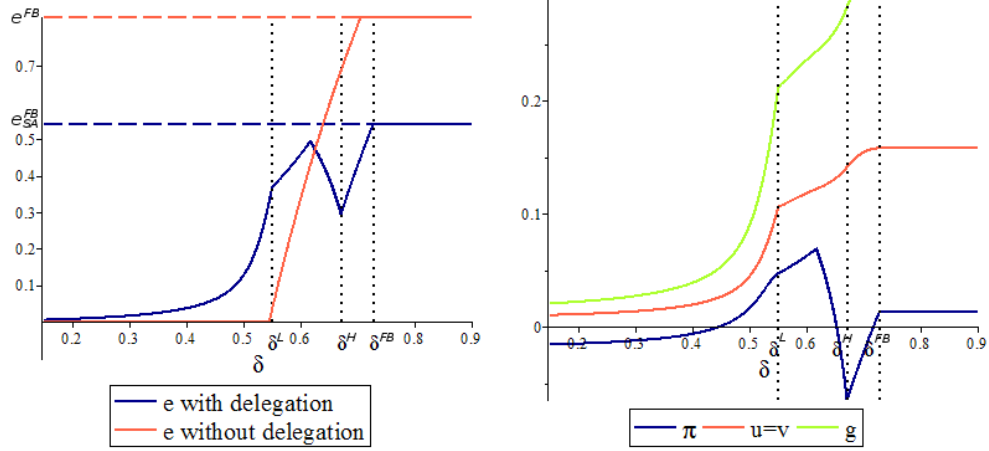
(a) Bonuses

(b) Bribes



(c) Effort

(d) Payoffs



$c(e) = .54 \times e^2$, $\overline{b} = 0.4, y = 0.9, \alpha = 0.6$, and $w = 0.05$

$\delta^H > \delta^L$, and there will be an intermediate surplus range where bonuses are used to induce effort.

3.2.2 Intermediate surplus

The players face a trade-off when deciding upon the bonus given when output is low, b_l. A higher b_l generates greater surplus directly, but it also decreases effort. Maximizing joint surplus gives us the following expression for b_l when $\overline{b} > b_l > 0$:

$$b_l = \frac{1-\alpha}{\alpha}(1-e)\,c''(e) - y + \frac{1}{\alpha}\frac{\delta g(e, b_h, b_l)}{1-\delta} \tag{2}$$

The first term of this expression stems from the direct gain in supervisor-agent surplus that an increase in b_l produces; the more likely low output is to occur, the higher this gain. The second term is the result of the reduction in expected output that an increase in b_l produces through the change in effort induced. The third term comes from the relational contracting constraint; higher surplus means that more effort can be induced through bribes, and hence b_l can be higher.

Since $(IC - DE)$ is binding, the effort level e is given by $c'(e) = \overline{b} - b_l + \frac{\delta g(e,\overline{b},b_l)}{1-\delta} - \alpha\overline{b}$. If we substitute in (2), we can see that effort is weakly decreasing in the discounted surplus if and only if $b_l > 0$. When $b_l > 0$, a decrease in the discounted surplus decreases b_l and hence the agent and supervisor have a greater incentive to increase effort. Instead, when $b_l = 0$, a lower surplus forces the supervisor to reduce the variation in bribes.

3.2.3 Low surplus

When the discounted surplus becomes low, i.e. $\delta = \delta^L$, the supervisor can only just promise to pay $b_h = \overline{b}$ and will not be able to combine this with any variation in bribes. When $\delta \leq \delta^L$, the bonus b_h will be the maximum that the supervisor can credibly promise, i.e. $b_h = \frac{1}{\alpha}\frac{\delta g(e,b_h,b_l)}{1-\delta}$, and b_l will again be either equal to b_h, zero or a solution to (2).

3.3 Discussion

An important implication of Proposition 1 is that there is sometimes, but not always, a trade-off between increasing production and reducing the surplus captured by intermediaries. The previous literature on relational contracts suggests that noncontractible production can be improved by increasing the discounted joint-surplus within relationships, for instance by increasing tenure or decreasing competitive pressure (Calzolari and Spagnolo, 2009; Board, 2011; Gibbons and Henderson, 2013). Yet those concerned with collusion argue that such policies will facilitate kickbacks (Martimort, 1999; Lambsdorff and Teksoz, 2005).

Our analysis implies that in some cases both effects may indeed occur simultaneously. Examples of such a trade-off can be found in public procurement, where in some instances policies designed to reduce corruption appear to have a negative impact on performance or output (Coviello, Guglielmo and Spagnolo, 2016; Lichand, Lopes and Medeiros, 2016). We also find, however, that in other cases there is no such trade-off, and decreasing discounted surplus will reduce collusion without any negative impacts. Examples of this can perhaps be found in the pubic procurement reforms studied by Palguta and Pertold (2015) and Lewis-Faupel et al. (2016), who find reducing discretion and decreasing interactions appears to reduce corruption with a non-negative effect on quality. In our model, these correspond to situations where the supervisor is on occasion dispensing the maximum possible bonus, but is never dispensing the minimum.

4 Optimal delegation for the principal

In the previous section, we ignored the role of the principal and treated the parameters w, α and $\overline{b}$ as exogenous. In this section, we consider that the principal sets these parameters at the beginning of the game. We first solve for the optimal parameters and describe the resulting contract. We then examine when the principal will prefer to delegate rather than undertake direct relational contracting.

4.1 How should the principal delegate?

We begin by writing the principal's payoff under delegation as a function of the supervisor-agent surplus g. Since the supervisor-agent contracts will be stationary, we can write the principal's payoff as follows:

$$\Pi = N\left[ey - c(e) - g - \underline{u}\right]$$

The principal effectively only cares about effort and the surplus given to the supervisor and agents; holding these constant, she is indifferent to the various potential compensation schemes. Furthermore, note from the definition of g that, when $\alpha < 1$, the principal can set g through the wage w.

The following proposition describes how the principal sets the parameters w, α and $\overline{b}$ to maximize her payoff. The principal sets supervisor-agent surplus such that the contract is of the 'low surplus' type described in Proposition 1. Moreover, she chooses α such that the supervisor's preferences are partly, but not entirely, aligned with her own, resulting in a contract with $b_h = \overline{b}$ and $b_l = 0$.

Proposition 2. *The supervisor-agent contract under optimal delegation will only use bonuses to induce effort; bonuses will be zero when output is low and at the cap when output is high. If the optimal delegated contract involves positive effort, then the optimal value of α for the principal lies strictly between 0 and 1. Effort will be strictly below that which maximizes total surplus, e^{FB}.*

The intuition behind the first part is that inducing effort through variation in bribes requires greater supervisor-agent surplus than using bonuses. The principal therefore prefers effort being induced using bonuses since it involves giving up less surplus. The principal can always improve on any contract that has variation in bribes by increasing the bonus cap and decreasing the wage in order to induce the same level of effort entirely through bonuses.

When it comes to setting α, the principal faces a trade-off. On the one hand, she wishes to decrease α in order to facilitate supervisor-agent relational contracting; easier relational contracting reduces the amount of

surplus the principal needs to give to the supervisor and agents. On the other hand, she needs to ensure that α is sufficiently high that the optimal supervisor-agent contract involves no bonuses when output is low, i.e. $b_l = 0$. A contract with $b_l > 0$ is not optimal since the principal could lower the bonus cap $\overline{b}$ and the wage w in order to induce a contract with the same variation in bonuses but with $b_l = 0$.

If the optimal contract involves positive effort, it must have $\alpha > 0$. Otherwise, if $\alpha = 0$, then the supervisor and agents will strictly prefer a contract that induces no effort, since effort is costly and brings them no rewards.

To examine the principal's trade-off, note that we have $s_l = s_h$ and $b_l = 0$, and hence $c'(e) = b_h = \frac{1}{\alpha}\frac{\delta g}{1-\delta}$. The principal therefore maximizes the following expression:

$$\Pi = N\left[ey - c(e) - \alpha\frac{1-\delta}{\delta}c'(e) - \underline{u}\right] \tag{3}$$

Since this expression is decreasing in α, the principal will set α at the lowest level at which it is possible to induce a contract with $b_l = 0$ and effort at the desired level. The principal must set α such that the supervisor and agent will prefer to induce effort rather than set $b_l = b_h$, which requires $\alpha \geq \frac{(1-e)c'(e)+c(e)}{(1-e)c'(e)+ye}$. She must also ensure that $b_l = 0$ is chosen rather than a level $0 < b_l < b_h$, but from (2) we will have no such interior solutions so long as $\alpha \geq \frac{(1-e)c''(e)}{(1-e)c''(e)+y-c'(e)}$.

The principal would therefore only set $\alpha = 1$ if she was giving the supervisor and agent sufficient surplus to induce first-best effort e^{FB}. This is not optimal for the principal, however, as at the margin inducing more effort involves giving up more surplus, and hence she will induce an effort level below e^{FB}. In other words, since the principal does not want to induce first-best effort, she does not require a supervisor who fully internalizes the benefits of effort.

Lemma 4. *As δ increases, it is optimal for the principal to increase the bonus cap and induce higher effort. In response to an increase in the agent's outside option, $\underline{u}$, the principal will increase the wage w but not α or the induced effort.*

If discounting is reduced, then relational contracting between the supervisor and agents eases and the marginal cost for the principal of inducing effort decreases. She therefore will chose to induce more effort. In the limit as $\delta \to 1$, the optimal value of α tends to 1 and the effort level tends to the principal's first-best, e^{FB}. On the other hand, it is clear from (3) that the optimal effort level (and hence the optimal α) for the principal does not change with respect to $\underline{u}$, since increasing the agent's outside option does not affect the marginal cost of inducing effort. It does, however, impact the payoff the principal receives from delegating, and hence may affect the principal's delegation decision, as we shall now explore.

4.2 When should the principal delegate?

The previous section considered the optimal way for the principal to delegate a relational contract to a supervisor. In some situations, delegation may be obliged; the leader of a government or large firm may simply be unable to manage all relevant relational contracts themselves. In other situations, the principal may have the choice between delegating the relational contract to a supervisor or managing it herself. In this section, we consider when such delegation may be in the principal's best interest.

The following proposition describes when the principal should delegate, assuming she does so optimally. If she does not delegate, we assume she undertakes direct relational contracting with the agents, achieving the results outlined in Section 3.1. The proposition states that there exists a range of discount factors for which delegating is strictly preferable and a higher range when direct relational contracting is strictly preferable.

Proposition 3. *There exist values $\underline{\delta}, \hat{\delta}$ and $\overline{\delta}$ with $0 \leq \underline{\delta} < \hat{\delta} \leq \overline{\delta} < 1$ such that:*

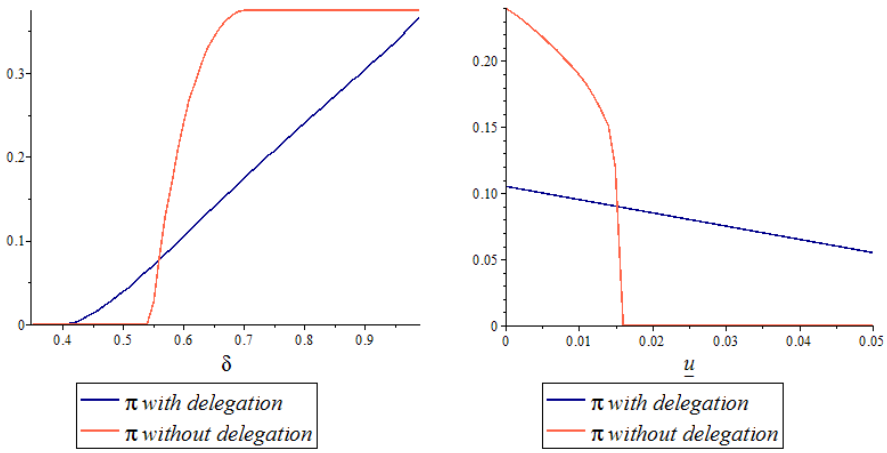
- *If $\delta > \overline{\delta}$, then the principal's payoff from optimally delegating is strictly below that from direct relational contracting.*
- *If $\hat{\delta} > \delta > \underline{\delta}$, then the principal's payoff from optimally delegating is strictly above that from direct relational contracting.*

If the discount factor is high, then relational contracting poses no problem, and there is no reason to delegate. The principal and agents can

implement first-best effort on their own, and the principal has no reason to share surplus with the supervisor. On the other hand, if the discount factor is low, then direct relational contracting is difficult and cannot sustain much effort. The principal would therefore prefer to delegate and thus generate more effort, since the extra surplus generated more than compensates for the part given to the supervisor.

Figure 3: Principal's payoffs with and without delegation

(a) Discount factor, δ (b) Agent's outside option, $\underline{u}$



$c(e) = .54 \times e^2$ and $y = 0.9$. When not plotted, $\delta = 0.55$ and $\underline{u} = 0$.

A similar logic applies for other variables affecting the potential discounted surplus, including the agents' outside option $\underline{u}$. Figure 3 demonstrates these results graphically by plotting the best payoffs that the principal can achieve with and without delegation when $c(e) = \frac{ce^2}{2}$. In this case we have a single crossing point in the payoffs, i.e. $\hat{\delta} = \overline{\delta}$, but this will not necessarily be the case for other functional forms.

4.3 Discussion

Proposition 2 tells us that the principal would like a supervisor whose payoffs are partly, but not completely, aligned with her own. The principal needs the supervisor to care somewhat about profit because otherwise no effort will be induced. This makes supervisor-agent relational contracting costly, which means that to get more effort the principal has to give up more surplus. As a result, the principal will not wish to induce first-best

effort when delegating, and hence there is no need to have a supervisor whose incentives align completely with her own. Instead, the principal would rather have a supervisor who cared less about profit in order to facilitate relational contracting.

An example of such behavior can perhaps be seen in the way in which Chinese firms deal with the practice of Guanxi, a system of informal relationships often formulated through gift exchange. Firms are well aware of the risks stemming from procurement and sales managers' personal relationships, since these can facilitate kickbacks and other malpractice (Millington, Eberhardt and Wilkinson, 2005). Yet, when it comes to hiring such personnel, Wiegel and Bamford (2015) find evidence that firms specifically hire people with personal Guanxi, and they cite the abiltiy of Guanxi to smooth inter-firm relationships as an important factor. Indeed, Schramm and Taube (2003) argue that whilst Guanxi networks facilitate corruption, this corruption itself helps to strengthen the legitimate transactions coordinated through Guanxi, in a way similar to that described in our model above.

Proposition 3 states that delegating can be beneficial for the principal when she has difficulty committing. An instance where this may be applicable is when governments delegate utility regulation to independent regulatory agencies. Empirical evidence suggests that such delegation has increased credibility in contexts where governments have difficulty committing to allowing firms a sufficient return on investments.[13] Amongst potential explanations, one is that independent regulators may be more easily 'captured' by the regulated firm, with Armstrong and Vickers (1996, p.303) noting in transition economies "a degree of capture might enhance the credibility of commitment to allow an adequate return on investment".

5 Alternative specifications and extensions

In this section we consider a number of alternative ways in which we could set up our model and discuss how this would affect our results. We begin by

[13]See Estache and Wren-Lewis (2009) for a survey. Evans, Levine and Trillas (2008) consider how delegation may solve this commitment problem and detail a number of cases.

considering how the principal would behave if α was exogenous, and then examine cases where she has more instruments at her disposal. Finally, we briefly sketch how the results would change were there to be a cost to side payments.

5.1 Exogenous α

The assumption that α is chosen by the principal may be reasonable in contexts where the principal designs the supervisor's contract, but unreasonable in others. For instance, Chassang and Padró i Miquel (2014) argue that payoff functions are rarely available as policy instruments in settings with corruption. An important situation where α may be exogenous is when it represents the supervisor's intrinsic motivations.

When α is high, the principal will induce an optimal 'low surplus' contract as described in Proposition 2. When α is low, however, such a contract cannot generate much effort, since the supervisor and agents would rather pay high bonuses when output is low than use bonuses to induce effort. In this case, the principal is better off allowing the supervisor no discretion (i.e. $\overline{b} = 0$) and delegation will decrease the principal's payoff. This is consistent with the results of Rasul and Rogger (2016), who find that subjective performance evaluation produces better results when supervisors are more intrinsically motivated.

5.2 Additional instruments for the principal

Proposition 2 showed that, under optimal delegation, bribes will occur and effort will be below the level that maximizes total surplus, e^{FB}. It is natural to ask whether these results would change if the principal had more instruments at her disposal, because instruments are likely to vary according to the institutional context. For instance, in the context of a government bureaucracy a principal may be unable to pay a supervisor a wage dependent on a relationship continuing, but in the private sector such compensation schemes may be possible. We therefore now consider the introduction of three new instruments: a wage for the supervisor, an initial transfer to the principal, and the ability to share output and costs by different proportions.

5.2.1 A wage for the supervisor

In the model we have used, only the agents receive a wage from the principal. Side payments therefore play an important role by allowing the supervisor and agents to split the surplus. Were the principal able to pay a wage to the supervisor, wages could be tailored so that the optimal contract involved no bribes in equilibrium. Delegation can therefore benefit the principal without collusion if instead the supervisor receives regular payments that are conditional on the relationship being maintained. Allowing for such a wage will not change any of the propositions above, since the principal would be indifferent between paying the supervisor directly and paying her through bribes. This may be one reason why kickbacks in the public sector receive more attention than kickbacks in the private sector - a greater number of instruments may allow private firms to replace the positive aspect of kickbacks and hence crack down on collusion.

One reason why delegated relational contracts may be sustained by side payments rather than payments from the principal is that the principal has limited information. To induce an optimal contract without collusion, the principal needs to know the relative bargaining powers of the supervisor and agent. Moreover, Hermalin (2015) argues that the principal may imperfectly observe when the intermediary is cooperating. If the supervisor might continue to receive payments from the principal after termination, then payments would also increase the supervisor's outside option. For this reason, side payments may be a more effective tool for sustaining delegated relational contracts.

The principal may also rule out direct transfers to the supervisor to reduce the potential for future renegotiations. Katz (1991) shows that if unobservable renegotiation is possible, then delegation loses much of its ability to solve commitment problems. Hence delegated relational contracting sustained by side payments may be more credible than delegation with relatively flexible principal-supervisor transfers.

5.2.2 A more general sharing rule

We previously assumed that the supervisor receives a share α of the profit. A more general contract would place a different weight on output to that

placed on bonuses. For instance, consider the following payoff function for the supervisor:

$$V_t = \sum_{i=1}^{i=N}(1-\delta)\mathbb{E}\left[\sum_{\tau=t}^{\infty}\delta^{\tau-t}d_{i\tau}\left[\alpha_Y y_{i\tau} - \alpha_b b_{i\tau} + w^S + S_{i\tau}\right]\right]$$

where α_Y and α_b are the weights placed on output and bonuses, and w_S is a per-agent wage paid by the principal to the supervisor. We assume that $\alpha_b \geq 0$ and $\alpha_Y \leq 1$.[14]

In this case, the nature of supervisor-agent optimal contracts do not change substantially, and Proposition 1 remains unchanged. The principal can take advantage of the more flexible contract and will not use a simple profit sharing rule of the type considered in the main model above. Nonetheless, the main result of Proposition 2 will still hold - the principal will induce effort lower than first best.

Proposition 4. *The principal can do better under a more general sharing rule, i.e. when $\alpha_Y \neq \alpha_b$, but will still induce effort lower than that which maximizes total surplus, e^{FB}.*

It is optimal for the principal to set $\alpha_Y = 1$ because a higher α_Y encourages the supervisor to induce more effort. There is no cost to the principal in increasing α_Y since she can extract surplus through the wage w^S. It is not optimal, however, for the principal to set $\alpha_b = 1$, just as it was not optimal for her to set $\alpha = 1$ in the simpler model. The principal will set α_b at the lowest level at which it is possible to induce a contract with $b_l = 0$ and effort at the desired level. This effort level will be below first-best following the same logic as behind Proposition 2 - greater effort requires giving more surplus to the supervisor. Hence she will set $\alpha_b < 1$ and induce an effort level $e < e^{FB}$.

Note that it may be optimal for the principal to set $\alpha_b = 0$. This makes relational contracting between the supervisor and agents unnecessary because the supervisor has no temptation not to pay promised bonuses. In

[14]If the principal could set $\alpha_Y > 1$, then she could achieve first best through delegation, but she would receive a negative share of profits each period. It is common in the literature to focus on nonnegative shares, as negative shares are difficult to implement in practice (Rayo, 2007).

this case, the principal only needs to give the supervisor and agents their outside options. The threat of supervisor misbehavior still exists, however, and hence the principal must limit the supervisor's discretion, $\overline{b}$, and thus the amount of effort she can motivate through bonuses.

5.2.3 Output-sharing with the agents

To keep the analysis simple, we have previously assumed that the supervisor received a share of total profit, but that agents do not receive any such explicit incentive. In some cases, however, it may be possible for agents to receive a profit share similar to that received by the supervisor. Within some firms, for instance, all employees receive profit shares, not just senior management (see Weitzman and Kruse, 1990, for a review). In this section we therefore extend our model to allow for this possibility.

We first consider how a supervisor would optimally share output with agents were she to have this tool at her disposal in addition to the relational contract. If there is only one agent, then clearly the supervisor can obtain her preferred effort level by passing all her output share to the agent. If there are multiple agents, however, Rayo (2007) shows that relational contracting may be optimal when it is impossible for each agent to receive the entire output. To see this in our model, consider the behavior of the supervisor and agents when the parameters set by the principal are given. Now, suppose that the supervisor can sell a share α_i to agent i, given the constraint that $\sum_i \alpha_i \leq \alpha_Y$, where α_Y is the share of output that the supervisor receives, as in Section 5.2.2. Then let $e_R(\delta)$ and $g_R(\delta)$ be the effort and supervisor-agent surplus generated at a given level of δ by the optimal supervisor-agent relational contract when there is no sharing of output, i.e. $\alpha_i = 0$. The following lemma then shows that, when N is large, it is optimal to share no output with the agents unless δ is small, and in particular smaller than the range of δ which Proposition 1 focuses on.

Lemma 5. *For a given N, if $\delta \geq \tilde{\delta}(N)$, where $\frac{\tilde{\delta}(N)}{1-\tilde{\delta}(N)} = \frac{1}{(N-1)e_R(\tilde{\delta}(N))}$, then it is optimal for the supervisor to share no output with the agents. Moreover, if $N > 1 + \frac{g_R(\delta^L)}{\overline{b}e_R(\delta^L)}$, then $\tilde{\delta}(N) < \delta^L$.*

The intuition behind this lemma is that sharing output with agents

has two opposing effects on the dynamic enforcement constraints. First, the enforcement constraints are effectively relaxed, since agents are motivated to exert effort with explicit non-discretionary incentives. Second, the enforcement constraints are tightened since the joint surplus in the relationship between a given agent and supervisor diminishes because part of the supervisor's output share has been given to the remaining agents. The size of the first effect is diminishing in N, while the size of the second effect is increasing in N. Hence, for a given discount factor δ, there is a critical number of agents N where above which giving any share of output to agents makes inducing effort more difficult.

Turning to the decisions of the principal, let us first consider the case where she delegates to the supervisor. It is clear that, if there is only one agent, the principal can achieve the first best. In particular, by setting $\alpha_Y = \alpha_b = 1$, the supervisor will have the same incentives as the principal. Then the supervisor will use only explicit incentives by sharing all the output with the agent. Since relational incentives are not used, the principal extract all the surplus via a negative w^S or w. For large N, however, delegation will involve relational contracting, and as in Proposition 4, it will be optimal to set $\alpha_b < 1$ to reduce the cost of this relational contracting."

If the principal has a choice over whether to delegate or not, the ability to share output with agents directly will make delegation less tempting. Indeed, if there is only one agent, there will be no need for a supervisor. If $N > 1 + \frac{1-\hat{\delta}}{\hat{\delta} e_R(\hat{\delta})}$, however, then $\tilde{\delta}(N) < \hat{\delta}$ and, since $e_R(\hat{\delta}) > 0$, there is always range of δ where the principal's best strategy is to delegate to a supervisor who will only use relational contracting.

5.2.4 Initial transfer to the principal

An important assumption in our model is that the wage w is fixed over time. Hence, when setting the wage, the principal faces a trade-off; a higher wage costs the principal directly, but it also increases the supervisor-agent discounted surplus and hence allows for greater effort. If the wage was allowed to vary over time, however, the principal would face no such trade-off. Instead, she could set high future wages to ensure a large supervisor-agent discounted surplus and a very low, potentially negative, initial wage

to extract this surplus ex ante.

In the extreme, allowing an initial fee to be paid to the principal by the supervisor or agent would allow the principal to achieve the first best. In particular, she could set $\alpha = 1$, extract all the surplus via an ex ante transfer, and then set future wages to be sufficiently large that the supervisor could credibly promise to pay bonuses of size y. But, if direct principal-agent relational contracting cannot also produce the first best, then achieving the first-best with delegation requires a wage larger than the total surplus generated each period. In some sense, therefore, the principal is improving her payoff not by delegation, but by being able to borrow in the first period and then invest in a financial product that only pays out if the relationship is sustained.

A realistic model allowing for initial transfers would therefore demand that wages (or total compensation) be capped at some level less than or equal to the period surplus. First-best effort would then only be achievable under delegation if it was achievable without delegation, since the supervisor will only be willing to induce e^{FB} if $\alpha = 1$. Overall, therefore, allowing wages to vary over time in a reasonable way will not result in first-best effort with delegation if the principal cannot achieve first best with direct relational contracting.

5.3 Costly side payments

We have assumed for simplicity that side payments between the supervisor and agent are costless except to the extent that they need to be self-enforced. In reality, however, side payments may be intrinsically costly. For instance, there may be a risk of punishment, and payments may be made in an inefficient way to avoid detection.

In this subsection we consider how our model would change if we make side payments costly. In particular, we assume that a side payment which costs an agent S only gives a benefit of κS to the supervisor, where $0 < \kappa \leq 1$. In order to keep the analysis simple, we make two additional assumptions. First, we assume that the supervisor has full bargaining power vis-a-vis each agent, meaning that the supervisor and agents will implement relational contracts which maximize v. Second, we assume that

the principal pays the supervisor a constant wage w^S, and that this wage must be set so that in equilibrium net bribes from the supervisor to the agents are non-negative. These two additional assumptions are relatively innocuous, but are made to avoid surplus being created in equilibrium through side payments from the supervisor to the agents.[15]

The optimal supervisor-agent contracts will not change significantly with these new assumptions - a monotonic transformation of the supervisor's payoff function tells us that introducing a cost of side payments κ is equivalent to the costless case where she receives a share of profit α/κ. Propositions 1-3 will therefore not change, but the principal will set α differently. In particular, the following proposition gives us comparative statics with respect to κ:

Proposition 5. *When side payments are more costly, i.e. κ is smaller, the principal will share less profit with the supervisor, delegate more often, and the optimal delegated contract will involve greater effort.*

When κ is smaller, the risk of collusion is reduced, and hence the principal can set a lower value of α and still avoid $b_l > 0$. Since α is lower, supervisor-agent relational contracting is easier, and the principal can transfer less surplus to the supervisor. Moreover, the lower value of α decreases the marginal cost to the principal of inducing effort, and hence she will ensure a higher level of output. Overall, lower κ increases the value of delegation to the principal. These results are consistent with Bloom, Sadun and Van Reenen (2012) and Bloom et al. (2013) who find that firms delegate more and increase production when there is either stronger rule of law or management practices which allow better monitoring of supervisors.

If side payments were impossible ($\kappa = 0$), then the principal could set $\alpha = 0$ and induce first-best effort. Thus in general the principal would prefer for collusion to be more costly so long as she has alternative instruments to reward the supervisor with.[16]

[15]We could alternatively assume that the side payment S was always positive, or that a cost was born by the agents for receiving bribes. Either assumption would lead to optimal contracts being potentially non-stationary, since an agent is limited in his ability to extract surplus from the supervisor and will therefore use a threat of lower production instead. The model would then share similarities with that of Fong and Li (2015), which can be seen as an example of the 'backloading' principal expounded by Ray (2002).

[16]Our results thus contrast with Strausz (1997), who in an alternative model of dele-

6 Conclusions

This paper has studied the impact of collusion on relational contracts, and in doing so has generated a number of empirical implications. On the one hand, we have seen that relationships that enable collusion may be the same as those that encourage valuable production. Hence the prevalence of corruption and kickbacks in contexts where explicit contract enforcement is weak may stem partly from the resulting reliance on relational contracting. Indeed, we have shown that principals may want their supervisors to be somewhat susceptible to collusion, and this may help to explain why politicians and firm owners frequently turn a blind eye to employees accepting kickbacks (Banfield, 1975). On the other hand, we have also shown that collusion can crowd out productive effort if the relationship between supervisor and agent is too strong. Indeed, the risks of collusion may be sufficiently large that a principal would rather manage a relationship herself than delegate to an intermediary. Governments and firms therefore face a delicate balancing act when it comes to making a trade-off between productive relational contracts and collusion.

An important next step will be to test the results of this paper in empirical work. Side payments are difficult to measure, but it should be possible to test our results on other variables such as output, discretionary rewards and discounted surplus. We could, for instance, test directly for a nonmonotonic relationship between effort and surplus in contexts where the principal is constrained in their ability to optimize, such as the public sector. In some circumstances it may also be possible to observe the extent to which supervisors use their discretion and test for the type of misuse predicted by the model. In other contexts, we may evaluate whether principals are behaving in a way consistent with the model by observing variation in delegation decisions and the incentives given to supervisors. A potentially under-explored area may be investigating firm owners' concerns with employee fraud in procurement, particularly in developing countries

gation finds outcomes are the same whether or not supervisor-agent collusion is possible (i.e. $\kappa = 1$ or $\kappa = 0$). We would obtain the same result if we made two changes to make our model closer to his. In particular, if we assumed that $e^{FB} = 1$ and the principal could pay the supervisor conditional on output and bonuses, then she could set $\alpha = 0$, offer the supervisor a large payment in exchange for setting $b_l = 0$, and achieve the first-best with output always high.

where explicit contracts are weak.

There are also multiple theoretical extensions to the model that would be valuable to pursue. For instance, we have assumed that the supervisor's preferences are known, but in reality there is uncertainty as to 'how corrupt' any individual is. Removing this assumption, possibly in a similar way to Chassang (2010) or Chassang and Padró i Miquel (2014), may reveal insights into how corruption and effort evolve over time. We may also ask whether collusive relational contracts make supervisors more likely to stick with the same firm over time. This may add an extra inefficiency, or help to enhance valuable loyalty.

Mathematical Appendix

Proof of Lemma 1. The proof is analogous that of Theorem 2 in Levin (2003), only we must define new side payments, rather than new bonus payments and wages, to produce the stationary contract. □

Proof of Lemma 2. For the first part, consider an optimal contract with $b_h < 0$. Then consider an alternative contract with bonus $b'_h = 0$ and side payment $s'_h = s_h - b_h$. It is simple to check that all the self-enforcing constraints are still satisfied. Moreover, this contract has a higher surplus, and therefore the original contract cannot be optimal. The same logic holds if $b_l < 0$.

For the second part, first suppose that $s_h > s_l$. Then, consider an alternative contract with $s'_l = s_h$, $b'_l = b_l + s_h - s_l$. This alternative contract must also be self-enforcing, yet surplus is greater. Hence the original contract is not optimal. In the case of bonuses, if $b_h < b_l$, then we can similarly consider an alternative contract with $b'_h = b_l$ and $s'(y) = s_h + b_l - b_h$. □

Proof of Lemma 3. First, consider an optimal contract with (DE_A) not binding. If $e < e^{FB}_{SA}$, then consider an alternative contract with $s'_l = s_l + \epsilon$. This contract induces higher effort and, for some $\epsilon > 0$, is self-enforcing. Thus we must have $e \geq e^{FB}_{SA}$. If $b_l < \overline{b}$, then consider a contract with $b'_l = b_l + \epsilon$ and $s'_l = s_l + \epsilon$. This contract generates higher surplus and, for some $\epsilon > 0$, is self-enforcing. Thus we must have $b_l = \overline{b}$. Lemma 2 then

implies $b_h = \overline{b}$.

Second, consider an optimal contract with (DE_S) not binding. If $e < e_{SA}^{FB}$, then consider an alternative contract with $s'_h = s_h - \epsilon$. This contract generates higher effort and, for some $\epsilon > 0$, will be self-enforcing. Hence we must have $e \geq e_{SA}^{FB}$. If $b_h < \overline{b}$, then we must have $s_l = s_h$, since otherwise we can construct an alternative self-enforcing contract that generates higher surplus with $b'_h = b_h + \epsilon$ and $s'_h = s_h + \epsilon$, for some $\epsilon > 0$. Since $e > 0$, it therefore follows that $b_l < b_h$, but now we can construct a self-enforcing contract with $b'_h = b_h + \epsilon$ and $b'_L = b_l + \epsilon$, for some $\epsilon > 0$. Hence we must have $b_h = \overline{b}$. Finally, if $b_l < \overline{b}$, then we can consider a contract with $b'_L = b_l + \epsilon$ and $s'_h = s_h - \epsilon$ (since $s_l \geq s_h$ from Lemma 2). But this contract is self-enforcing for some $\epsilon > 0$ and has higher surplus. Hence we must have $b_l = \overline{b}$.

Therefore, if either (DE_A) or (DE_S) is not binding, we must have $b_h = b_l = \overline{b}$ and $e \geq e_{SA}^{FB}$. Summing (DE_A) and (DE_S) and substituting into (IC) gives $\alpha\overline{b} + c'(e) < \frac{\delta}{1-\delta}(v + u - \underline{u}) \leq \frac{\delta g(e_{SA}^{FB}, \overline{b}, \overline{b})}{1-\delta}$. But, since $e \geq e_{SA}^{FB}$, we must have $c'(e) \geq \alpha y$, which implies $\delta \geq \delta^{FB}$. □

Proof of Proposition 1. If $\delta \geq \delta^{FB}$, then the first-best contract is self-enforcing. This contract is 'high surplus' in the sense of the proposition. For the rest of the proof we consider the case when $\delta < \delta^{FB}$. We first consider how the characteristics of the optimal contact change as a function of δ, and then how effort e changes as a function of δ for each contract type.

First, note that both the surplus and effort level are increasing in b_h, and hence $b_h = \min\{\overline{b}, \frac{1}{\alpha}\frac{\delta g^*}{1-\delta}\}$, where g^* is the maximum surplus generated by any self-enforcing contract. From Lemma 3, $\overline{b} \geq \frac{1}{\alpha}\frac{\delta g^*}{1-\delta}$ implies $s_h = s_l$ and hence we can define δ^L such that $\overline{b} = \frac{\delta^L g^*}{1-\delta^L}$. There is a unique value of δ that solves this equation since g^* is weakly increasing in δ. We now consider in turn the cases when $\delta \geq \delta^L$ and $\delta \leq \delta^L$.

If $\delta \geq \delta^L$, then from Lemma 3 we have $b_l = \frac{\delta g(e, \overline{b}, b_l)}{1-\delta} + (1-\alpha)\overline{b} - c'(e)$. The supervisor and agent's problem is to maximize the surplus function $g_1(e) = g\left(e, \overline{b}, b_l\right)$ with respect to e subject to the boundary conditions on

b_l, i.e. $0 \leq b_l \leq \overline{b}$. Differentiating this function with respect to e gives us:

$$g_1'(e) = \frac{\alpha y + (1-\alpha)\left(\alpha\overline{b} - (1-e)c''(e)\right) - \alpha c'(e) - \frac{\delta g_1(e)}{1-\delta}}{1-\delta-\delta(1-e)(1-\alpha)} \tag{4}$$

Let $\overline{g}_1$ and $\underline{g}_1$ be the surpluses at the upper and lower boundaries and let $\tilde{g}_1$ be the surplus at the maximal interior solution. Define effort levels $\overline{e}_1$, $\underline{e}_1$ and $\tilde{e}_1$ as the effort induced at these three potential solutions. Since we wish to characterize the optimal contract as a function of δ, we differentiate each of these surpluses with respect to δ, giving the following equations:

$$\begin{aligned}
\frac{d\overline{g}_1}{d\delta} &= \frac{\overline{g}_1}{1-\delta}\frac{\alpha y - c'(\overline{e}_1)}{(1-\delta)c''(\overline{e}_1) - \delta(\alpha y - c'(\overline{e}_1))} \\
\frac{d\tilde{g}_1}{d\delta} &= \frac{\tilde{g}_1}{1-\delta}\frac{(1-\alpha)(1-\tilde{e}_1)}{1-\delta-(1-\alpha)(1-\tilde{e}_1)\delta} \\
\frac{d\underline{g}_1}{d\delta} &= \frac{\underline{g}_1}{1-\delta}\frac{\alpha y + (1-\alpha)\overline{b} - c'(\underline{e}_1)}{(1-\delta)c''(\underline{e}_1) - \delta(\alpha y + (1-\alpha)\overline{b} - c'(\underline{e}_1))}
\end{aligned}$$

From (4), at any interior solution we have $c'(\tilde{e}_1) = y - \frac{1-\alpha}{\alpha}(1-\tilde{e}_1)c''(\tilde{e}_1) - \frac{1-\alpha}{\alpha}\left(\frac{\delta\tilde{g}_1}{1-\delta} - \alpha\overline{b}\right)$. Thus, at any δ where $\overline{g}_1 = \tilde{g}_1$, we have we have $c'(\overline{e}_1) < c'(\tilde{e}_1)$ and hence $\frac{d\overline{g}_1}{d\delta} > \frac{d\tilde{g}_1}{d\delta}$. There therefore exists a single value of δ such that for all higher values the upper boundary is preferable to an interior solution, and for all lower values the interior solution is preferable. We can show similarly that, at any δ where $\overline{g}_1 = \underline{g}_1$, we have $\frac{d\overline{g}_1}{d\delta} > \frac{d\underline{g}_1}{d\delta}$. Hence there exists a value δ_1^H such that the optimal solution has $b_l = \overline{b}$ if and only if $\delta \geq \delta_1^H$.

If $\delta \leq \delta^L$, then the supervisor and agent wish to maximize the surplus function $g_2(e) = g\left(e, \frac{1}{\alpha}\frac{\delta g_2(e)}{1-\delta}, \frac{1}{\alpha}\frac{\delta g_2(e)}{1-\delta} - c'(e)\right)$ with respect to e subject to the boundary conditions that $0 \leq b_l \leq \frac{1}{\alpha}\frac{\delta g_2(e)}{1-\delta}$. A similar analysis then gives us that there exist a value δ_2^H such that the optimal solution has $b_l = b_h$ if and only if $\delta \geq \delta_2^H$.

To define δ^H, note that when when $\delta = \delta^L$ both cases apply, and hence it cannot be that $\delta_2^H \leq \delta^L < \delta_1^H$. We therefore define $\delta^H = \min\{\delta_1^H, \delta_2^H\}$ and this then defines the high surplus case as stated in the proposition.

Finally, let us consider the relationship between e and δ. For high surplus contracts, e is given by Lemma 3, i.e. $\frac{c'(e)+\alpha\bar{b}}{g(e,\bar{b},\bar{b})} = \frac{\delta}{1-\delta}$. Differentiating the LHS with respect to e gives us that e is increasing in δ for high sur-

plus contracts. Similar differentiations for contracts with low surplus, or intermediate surplus with $b_l = 0$, show that e is increasing in δ for these cases.

For intermediate surplus optimal contracts with $b_l > 0$, we have $e = \tilde{e}_1$ where $g_1'(\tilde{e}_1) = 0$ and $g_1''(\tilde{e}_1) < 0$. Differentiating (4) by e gives:

$$g_1''(\tilde{e}_1) = (1-\delta)\frac{(1-2\alpha)c''(\tilde{e}_1) - (1-\alpha)(1-\tilde{e}_1)c'''(\tilde{e}_1)}{1-\delta-(1-\alpha)\delta(1-\tilde{e}_1)}$$

Note that for this contract to be optimal we must have $1-\delta-(1-\alpha)\delta(1-\tilde{e}_1) > 0$, since otherwise increasing b_l and s_l simultaneously relaxes $(IC-DE)$. $g_1''(\tilde{e}_1) < 0$ therefore implies $(1-2\alpha)c''(\tilde{e}_1)-(1-\alpha)(1-\tilde{e}_1)c'''(\tilde{e}_1) < 0$. We then differentiate $g_1'(\tilde{e}_1) = 0$ implicitly by δ to obtain:

$$\frac{d\tilde{e}_1}{d\delta} = \frac{g_1(\tilde{e}_1)}{(1-\delta)^2\left((1-2\alpha)c''(\tilde{e}_1) - (1-\alpha)(1-\tilde{e}_1)c'''(\tilde{e}_1)\right)}$$

This expression is negative in any optimal intermediate contract with $b_l > 0$, and this thus completes the proof. □

Proof of Proposition 2. We begin by considering how the principal should set parameters if she wishes to induce a contract with $s_l = s_h$. We then show that inducing such a contract is indeed optimal, and finally that this implies $e < e^{FB}$.

The principal can induce a contract with $s_l = s_h$ by setting $\overline{b} = \frac{1}{\alpha}\frac{\delta g}{1-\delta}$. Then, from Lemma 3, we have $c'(e) = \frac{1}{\alpha}\frac{\delta g}{1-\delta} - b_l$, where b_l either takes the value 0, $\overline{b}$ or a solution to (2). Note that $\frac{d^2 g_2(e)}{d\alpha de} \geq 0$ when $g_2'(e) = 0$, and hence there exists a value $\hat{\alpha}$ such that the optimal contract will have $b_l = 0$ if and only if $\alpha \geq \hat{\alpha}$. Moreover, $\hat{\alpha} > 0$ since at $\alpha = 0$ the supervisor and agent would clearly prefer setting $b_l = b_h$. If parameters are such that $b_l > 0$, then the principal would prefer to increase α since effort is increasing in α. Hence the principal will always set $\alpha \geq \hat{\alpha}$ and we then have $c'(e) = \frac{1}{\alpha}\frac{\delta g}{1-\delta}$.

Now, suppose that parameters induce a supervisor-agent contract with $s_l > s_h$. Then, from the proof of Proposition 1, we have $b_h = \overline{b}$. We will consider the three possibilities for b_l, and show that in each case the principal can adjust parameters to improve her payoff. First, suppose that

$b_l = 0$. Then from Lemma 3 we have $c'(e) = \frac{\delta g}{1-\delta} + (1-\alpha)\overline{b}$ and hence by increasing $\overline{b}$ and adjusting w to keep g constant the principal can increase effort without giving up extra surplus. Second, suppose that $0 < b_l < \overline{b}$. In this case, Proposition 1 has shown that effort is decreasing in the surplus level, and hence the principal can do better by decreasing g. Third, suppose that $b_l = b_h = \overline{b}$. If $\overline{b} > 0$, the principal can increase effort while holding surplus constant through reducing $\overline{b}$. If $\overline{b} = 0$, then $c'(e) = \min\left\{\alpha y, \frac{\delta g}{1-\delta}\right\}$. If $g > \frac{(1-\delta)\alpha y}{\delta}$, then the principal can do better by reducing g. If $g \leq \frac{(1-\delta)\alpha y}{\delta}$, then $c'(e) = \frac{\delta}{1-\delta} g$, which for any given g is equal or less than the effort achieved in the optimal contract with $s_l = s_h$. Hence the principal can always weakly improve on any contract with $s_l > s_h$.

The principal therefore effectively sets α and e to maximize Π subject to the condition that $\alpha \geq \hat{\alpha}$. From (3), Π is decreasing in α and hence the principal will set $\alpha = \hat{\alpha} = \max\left\{\frac{(1-e)c'(e)+c(e)}{(1-e)c'(e)+ye}, \frac{(1-e)c''(e)}{(1-e)c''(e)+y-c'(e)}\right\}$. We then have $\frac{d\hat{\alpha}}{de}\Big|_{e=e^{FB}} = \frac{1}{1-e^{FB}}$, and hence $\Pi'(e^{FB}) < 0$. It is therefore not optimal for the principal to set $e = e^{FB}$ and hence the principal will set $\alpha < 1$. □

Proof of Lemma 4. Let e^* be the level of effort that maximizes (3). We then have $\Pi'(e^*) = y - c'(e^*) - \frac{1-\delta}{\delta}\alpha(e^*)c''(e^*) - \frac{1-\delta}{\delta}\alpha'(e^*)c'(e^*) = 0$. Differentiating by δ gives us $\frac{de^*}{d\delta} = -\frac{\alpha(e^*)c''(e^*)+\gamma\alpha'(e^*)c'(e^*)}{\delta^2\Pi''(e^*)} = -\frac{y-c'(e^*)}{\delta(1-\delta)\Pi''(e^*)}$, which is positive since $e^* < e^{FB}$ from Proposition 2. □

Proof of Proposition 3. The existence of $\overline{\delta} \in (0,1)$ is straightforward. Relational contracting is only a constraint in the no-delegation case when (1) does not hold at e^{FB}. When e^{FB} is achievable without delegation, it is better for the principal not to delegate since delegation involves ceding surplus.

To show the existence of $\hat{\delta} \in (0, \overline{\delta})$, note that positive effort without delegation requires a value $e > 0$ that satisfies $ey - c(e) - \frac{1-\delta}{\delta}c'(e) \geq \underline{u}$. Define $\overline{e}(\delta)$ as the value of e that maximizes the LHS of this inequality for a given value of δ. Then, since the LHS of this inequality is increasing in δ, we can define $\hat{\delta} \in (0, \overline{\delta})$ to be such that $\overline{e}(\hat{\delta})y - c(\overline{e}(\hat{\delta})) - \frac{1-\hat{\delta}}{\hat{\delta}}c'(\overline{e}(\hat{\delta})) = \underline{u}$. In other words, $\hat{\delta}$ is the lowest value of δ such that there exists a relational contract sustaining positive effort without delegation.

For the principal to receive a positive payoff with delegation, we re-

quire values of e and α that lead to a positive value of the expression $ey - c(e) - \alpha\frac{1-\delta}{\delta}c'(e) - \underline{u}$. With an appropriate value of $\alpha < 1$, the principal can induce effort $e = \overline{e}(\hat{\delta})$ at $\delta = \hat{\delta} - \epsilon$ to obtain a payoff of $c'(\overline{e}(\hat{\delta}))\left(\frac{1-\hat{\delta}}{\hat{\delta}} - \alpha\frac{1-\hat{\delta}-\epsilon}{\hat{\delta}-\epsilon}\right)$. Hence, for $\epsilon > 0$ sufficiently small, the principal can achieve a payoff greater than her outside option under delegation when $\delta = \hat{\delta} - \epsilon$. Since the principal cannot achieve a payoff greater than her outside option without delegation at this level of δ, delegation is strictly preferable. □

Proof of Proposition 4. The proof that the principal will induce a 'low surplus' contract is as for Proposition 2, only now the contract introduces effort such that $c'(e) = \frac{1}{\alpha_b}\frac{\delta g}{1-\delta}$. The corresponding constraints on α_b are derived as before, and require that $\alpha_b \geq \max\left\{\frac{(1-e)c'(e)+c(e)-ye}{(1-e)c'(e)}, \frac{(1-e)c''(e)-y}{(1-e)c''(e)-c'(e)}\right\}$. It is straightforward to see that these constraints allow a lower level of α_b for any given effort level, and hence the principal will be required to transfer less surplus to induce any effort level e. We can also note that $e = e^{FB}$ implies $c'(e) = y$ and hence requires $\alpha_b \geq 1$. Since the principal is maximizing $\Pi = N\left[ey - c(e) - \alpha_b\frac{1-\delta}{\delta}c'(e) - \underline{u}\right]$, such an effort cannot be optimal. □

Proof of Lemma 5. Consider the optimal supervisor-agent contract that occurs without any sharing of output with the agents. If the dynamic enforcement constraints are not binding, then relational contracting is achieving the maximum surplus and hence there is no benefit to sharing output. If the dynamic enforcement constraints are binding, then sharing output with agents may increase surplus if doing so increases agents' effort. Since agents are identical, there is no reason to give them different shares of output, and we therefore consider that each agent receives $\frac{\alpha_A}{N}Y$, where $\alpha_A \leq \alpha_Y$. The effort of agent i will then be determined by $c'(e_i) = \frac{\alpha_A}{N}y + \frac{\delta g(e_i)}{1-\delta}$, where $g(e_i)$ is the expected joint surplus in the relationship between the supervisor and agent i, i.e. $g(e_i) = e_i y\left(\alpha_Y - \alpha_A + \frac{\alpha_N}{N}\right) + (1-\alpha_b)(e_i b_h + (1-e_i)b_l) - c(e_i) + w + \frac{w^S}{N} - \underline{u}$.

Differentiating by α_A gives $\left(c''(e_i) - \frac{\delta g'(e_i)}{1-\delta}\right)\frac{de_i}{d\alpha_A} = y\left(\frac{1}{N} - e_i\frac{\delta}{1-\delta}\frac{N-1}{N}\right)$. Since the dynamic enforcement constraint is binding, we must have $c''(e_i) > \frac{\delta g'(e_i)}{1-\delta}$. Hence if $\delta \geq \tilde{\delta}(N)$, then $\frac{de_i}{d\alpha_A} < 0$ for all $e_i \leq e_R(\tilde{\delta}(N))$, and thus effort is maximized by setting $\alpha_A = 0$. Then, from the proof of Proposi-

tion 1, δ^L is defined by $\frac{\delta^L}{1-\delta^L} = \frac{\bar{b}}{g_R(\delta^L)}$. Hence $\tilde{\delta}(N) < \delta^L$ is equivalent to $\frac{\bar{b}}{g^*} > \frac{1}{(N-1)e_R(\delta^L)}$. □

Proof of Proposition 5. The proof that the principal will induce a 'low surplus' contract is as for Proposition 2. In order to avoid waste, the principal will set the combination of w and w^S to ensure that the optimal contract involves no side payments, and hence she will maximize (3) as before. The constraints on α now become constraints on α/κ and, since α is chosen to make on of these constraints binding, the optimal α will be lower. It is then clear from (3) that, even for a given e, the principal's payoff under delegation is increasing in κ, and hence it is straightforward to see that delegation is more attractive. Moreover, from (3), we can see that a lower α means the principal will induce an effort e closer to first best. □

References

Antràs, Pol, and C. Fritz Foley. 2015. "Poultry in Motion: A Study of International Trade Finance Practices." *Journal of Political Economy*, 123(4): pp. 853–901.

Armstrong, Mark, and John Vickers. 1996. "Regulatory reform in telecommunications in Central and Eastern Europe." *Economics of Transition*, 4(2): 295–318.

Banfield, Edward C. 1975. "Corruption as a Feature of Governmental Organization." *The Journal of Law & Economics*, 18(3): 587–605.

Blader, Steven, Claudine Gartenberg, Rebecca Henderson, and Andrea Prat. 2015. "The Real Effects of Relational Contracts." *American Economic Review*, 105(5): 452–56.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. 2013. "Does Management Matter? Evidence From India." *The Quarterly Journal of Economics*, 128(1): 1–51.

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen. 2012. "The Organization of Firms Across Countries." *Quarterly Journal of Economics*, 127(4): 1663–1705.

Board, Simon. 2011. "Relational Contracts and the Value of Loyalty." *American Economic Review*, 101(7): 3349–3367.

Calzolari, Giacomo, and Giancarlo Spagnolo. 2009. "Relational contracts and competitive screening."

Chassang, Sylvain. 2010. "Building Routines: Learning, Cooperation, and the Dynamics of Incomplete Relational Contracts." *American Economic Review*, 100(1): 448–465.

Chassang, Sylvain, and Gerard Padró i Miquel. 2014. "Corruption, Intimidation, and Whistle-blowing: a Theory of Inference from Unverifiable Reports." National Bureau of Economic Research NBER Working Paper 20315.

Cole, Shawn, and Anh Tran. 2011. "Evidence from the Firm: A New Approach to Understanding Corruption." In *International Handbook on the Economics of Corruption Vol. II.* , ed. Susan Rose-Ackerman and Tina Soriede, 408–427. Edward Elgar Publishing.

Coviello, Decio, Andrea Guglielmo, and Giancarlo Spagnolo. 2016. "The effect of discretion on procurement performance." *Management Science*.

Dufwenberg, Martin, and Giancarlo Spagnolo. 2015. "Legalizing bribe giving." *Economic Inquiry*, 53(2): 836–853.

Estache, Antonio, and Liam Wren-Lewis. 2009. "Toward a Theory of Regulation for Developing Countries: Following Jean-Jacques Laffont's Lead." *Journal of Economic Literature*, 47: 729–770.

Evans, Joanne, Paul Levine, and Francesc Trillas. 2008. "Lobbies, Delegation and the Under-Investment Problem in Regulation." *International Journal of Industrial Organization*, 26(1): 17–40.

Fairburn, James A, and James M Malcomson. 2001. "Performance, Promotion, and the Peter Principle." *Review of Economic Studies*, 68(1): 45–66.

Fong, Yuk-fai, and Jin Li. 2015. "Relational Contracts, Limited Liability, and Employment Dynamics." Mimeo.

Fong, Yuk-Fai, and Jin Li. 2016. "Information Revelation in Relational Contracts." *The Review of Economic Studies*.

Francois, Patrick, and Joanne Roberts. 2003. "Contracting Productivity Growth." *Review of Economic Studies*, 70(1): 59–85.

Gibbons, Robert, and Rebecca Henderson. 2013. "What Do Managers Do? Exploring Persistent Performance Differences Among Seemingly Similar Enterprises." In *The Handbook of Organizational Economics*. , ed. Robert Gibbons and John Roberts, 680–731. Princeton University Press.

Goldlücke, Susanne, and Sebastian Kranz. 2013. "Renegotiation-Proof Relational Contracts." *Games and Economic Behavior*, 80: 157–178.

Hermalin, Benjamin E. 2015. "Why Whine about Wining and Dining?" *Journal of Law, Economics, and Organization*.

Katz, Michael L. 1991. "Game-playing Agents: Unobservable Contracts as Precommitments." *RAND Journal of Economics*, 22(3): 307–328.

Khanna, Vikramaditya, E. Han Kim, and Yao Lu. 2015. "CEO Connectedness and Corporate Fraud." *The Journal of Finance*, 70(3): 1203–1252.

Kofman, Fred, and Jacques Lawarree. 1993. "Collusion in hierarchical agency." *Econometrica: Journal of the Econometric Society*, 629–656.

Lambsdorff, Johann Graf, and Sitki Utku Teksoz. 2005. "Corrupt Relational Contracting." In *The New Institutional Economics of Corruption*. , ed. Johann Graf Lambsdorff, Markus Taube and Matthias Schramm, 154–168. Routledge.

Levin, Jonathan. 2002. "Multilateral Contracting and the Employment Relationship." *Quarterly Journal of Economics*, 1075–1103.

Levin, Jonathan. 2003. "Relational Incentive Contracts." *American Economic Review*, 93(3): 835–857.

Lewis-Faupel, Sean, Yusuf Neggers, Benjamin A Olken, and Rohini Pande. 2016. "Can Electronic Procurement Improve Infrastructure Provision? Evidence from Public Works in India and Indonesia." *American Economic Journal: Economic Policy*.

Lichand, Guilherme, Marcos FM Lopes, and Marcelo C Medeiros. 2016. "Is Corruption Good For Your Health?" *Mimeo*.

Macchiavello, Rocco, and Ameet Morjaria. 2015. "The Value of Relationships: Evidence from a Supply Shock to Kenyan Rose Exports." *American Economic Review*, 105(9): 2911–45.

Malcomson, James M. 2013. "Relational Incentive Contracts." In *The Handbook of Organizational Economics*. , ed. Robert Gibbons and John Roberts, 1014–1065. Princeton University Press.

Martimort, David. 1999. "The Life Cycle of Regulatory Agencies: Dynamic Capture and Transaction Costs." *The Review of Economic Studies*, 66(4): 929–947.

Martimort, David, and Thierry Verdier. 2004. "The Agency Cost of Internal Collusion and Schumpeterian growth." *Review of Economic Studies*, 71(4): 1119–1141.

Milgrom, Paul R. 1988. "Employment Contracts, Influence Activities, and Efficient Organization Design." *Journal of Political Economy*, 96(1): 42–60.

Miller, David A, and Joel Watson. 2013. "A Theory of Disagreement in Repeated Games with Bargaining." *Econometrica*, 81(6): 2303–2350.

Millington, Andrew, Markus Eberhardt, and Barry Wilkinson. 2005. "Gift giving, guanxi and illicit payments in buyer–supplier rela-

tions in China: Analysing the experience of UK companies." *Journal of Business Ethics*, 57(3): 255–268.

Nkamleu, Guy Blaise, and Bernadette Dia Kamgnia. 2014. "Uses and Abuses of Per-diems in Africa: A Political Economy of Travel Allowances." African Development Bank Group, AfDB Working Paper 196.

Olsen, Trond E, and Gaute Torsvik. 1998. "Collusion and Renegotiation in Hierarchies: A Case of Beneficial Corruption." *International Economic Review*, 413–438.

Palguta, Ján, and Filip Pertold. 2015. "Manipulation and Active Waste in Public Procurement: Evidence from the Introduction of Discretionary Thresholds." *mimeo.*

Rasul, Imran, and Daniel Rogger. 2016. "Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service." *Economic Journal*, Forthcoming.

Ray, Debraj. 2002. "The time structure of self-enforcing agreements." *Econometrica*, 70(2): 547–582.

Rayo, Luis. 2007. "Relational Incentives and Moral Hazard in Teams." *Review of Economic Studies*, 74(3): 937–963.

Schramm, Matthias, and Markus Taube. 2003. "The Institutional Economics of legal institutions, Guanxi, and Corruption in the PR China." In *Fighting Corruption in Asia: Causes, Effects, and Remedies.* 271–96. World Scientific.

Spagnolo, Giancarlo. 2005. "Managerial incentives and collusive behavior." *European Economic Review*, 49(6): 1501–1523.

Spagnolo, Giancarlo. 2012. "Reputation, Competition, and Entry in Procurement." *International Journal of Industrial Organization*, 30(3): 291 – 296.

Strausz, Roland. 1997. "Delegation of Monitoring in a Principal-Agent Relationship." *Review of Economic Studies*, 64(3): 337–357.

Thiele, Veikko. 2013. "Subjective Performance Evaluations, Collusion, and Organizational Design." *Journal of Law, Economics, and Organization*, 29(1): 35–59.

Tirole, Jean. 1986. "Hierarchies and Bureaucracies: On the Role of Collusion in Organizations." *Journal of Law, Economics, and Organization*, 2(2): 181–214.

Vickers, John. 1985. "Delegation and the Theory of the Firm." *The Economic Journal*, 95: 138–147.

Weitzman, Martin, and Douglas Kruse. 1990. "Profit Sharing and Productivity." In *The Econoimc Nature of the Firm.* , ed. Louis Putterman and Randall S. Kroszner, 288–296. Cambridge University Press.

Wiegel, Wladimir, and David Bamford. 2015. "The role of guanxi in buyer–supplier relationships in Chinese small-and medium-sized enterprises–a resource-based perspective." *Production Planning & Control*, 26(4): 308–327.