Relational Contracts and the Value of Loyalty

SIMON BOARD*

This Version: August 12, 2009
First Version: July 2006.

Abstract

This paper characterises the optimal contract for a principal who repeatedly chooses among a set of potential trading partners (agents) under the threat of holdup. As the economy evolves, the principal would like to trade with different agents; however, the possibility of ex–post opportunism allows agents to collect rents and creates a fixed cost of initiating a new relationship. In the optimal contract, the principal divides agents into "insiders", with whom she has previously traded, and "outsiders", with whom she has never traded. The principal then uses insiders efficiently, while being biased against outsiders. The optimal contract is self–enforcing if the principal is sufficiently patient. It can also be implemented by an "employment contract" that is robust to asymmetric information. The model thereby provides a theory of the firm, where both firm size and growth are endogenous.

1 Introduction

Holdup, whereby investment is deterred by the threat of ex–post opportunism, is one of the most pervasive forces in economics. In early trading communities, merchants who traded in foreign markets worried that part of their shipments would be embezzled by their agents (Greif (1993)). In modern–day Vietnam, where contracts enjoy little legal protection, 25% of suppliers have had a customer fail to pay for a product after it has been delivered (McMillan and Woodruff (1999)). And in developed countries, despite the presence of a functioning legal system, the majority of transactions are covered by no contract, by contradictory contracts or by contracts that are not legally enforceable (Macaulay (1963)).

In all these cases, ongoing relationships are used to mitigate the holdup problem. However, maintaining a relationship can reduce of scope of trade. The purpose of this paper is to make

*Department of Economics, UCLA. http://www.econ.ucla.edu/sboard/. I am grateful to Li Hao, Christian Hellwig, Hugo Hopenhayn, Jon Levin, Igor Livshits, Bentley MacLeod, Niko Matouschek, Bill Rogerson, Andy Skrzypacz and Bill Zame for many helpful comments. I also thank seminar audiences at CEMFI, CETC (Montreal), HBS, IIOC (Savannah), NASM (Duke), Northwestern, UBC, UC Berkeley, UCLA and Western Ontario. JEL codes: C73, L14. Keywords: Relational contracts, Repeated games, Loyalty, Holdup.

this tradeoff precise. In doing so, we derive the optimal relational contract, show that it induces loyalty, and characterise resulting the distortions. The model enables us to make predictions about when a firm will switch suppliers, and how these switches depend on the length of the relationship, the number of suppliers, and the shocks to the economy. The following examples illustrate the practical importance of loyalty in overcoming the holdup problem.

In 1984, the U.S. government started awarding its \$200bn worth of contracts through "full and open competition", either in the form of sealed bids or competitive negotiations. In order to reduce corruption, the evaluation panel was instructed to ignore subjective information, such as prior performance. Kelman (1990) examines computer procurement, comparing the government's policy with the purchasing behaviour of private firms. First, he finds that private firms were considerably more loyal to their suppliers. In an extensive survey, 58% of government contracts were awarded to the incumbent, compared to 78% for private firms. Moreover, 65% of government contracts were awarded to the lowest bidder, compared to 41% for private contracts. Second, Kelman finds that private firms were considerably more satisfied with their contracts. The government contracts rated particularly badly on 'keeping promises' and 'sticking to the contracted delivery schedule'. For example, Kelman reports that vendors held up the government by withdrawing key personnel, investing in fewer on–site personnel and, perhaps most importantly, providing poor advice and few creative ideas.

The importance of loyalty is also illustrated by the 1980s automobile industry.[1] At this time, General Motors awarded one–year contracts for each component through competitive bidding. While GM exploited the static gains from trade, it had concerns about quality, necessitating inspections of deliveries. In contrast, Toyota was more loyal, automatically renewing contracts for the life of the model and, for new models, favouring preferred bidders and suppliers with high performance ratings. As a result, Toyota overcame the holdup problem, trusting suppliers to meet the required quality standards, and installing parts without inspection. While it is not ex–ante obvious which system is better, it is notable that Toyota outsourced significantly more production, and that GM later tried, and failed, to adopt the Toyota model.

Finally, Johnson, McMillan, and Woodruff (2002) analyse the incidence of loyalty using survey of transitional Eastern European countries. When asked whether they would change suppliers for a 10% price discount, over half the firms surveyed said they would pass up on the new deal, in whole or in part. Johnson et al. find that these switching costs were significantly larger in countries with less effective court systems and for goods that are custom–built, suggesting that loyalty is a consequence of the relational contracts used to temper opportunistic behaviour.

Together, these examples exhibit two key features. First, when given the choice, firms are loyal to their suppliers, rather than simply maximising the static gains from trade. Second,

[1]See Asanuma (1989) and Milgrom and Roberts (1997).

loyalty mitigates the holdup problem, and is used more frequently when contracts are hard to write. This paper will present an explanation of these observations; it will also derive the optimal contract in such an environment.

1.1 Outline of the Paper

We start with a basic holdup model. Each period, a principal invests in one of N agents.[2] The winning agent then has the opportunity to hold up the principal. We suppose the cost of investment differs across time and across agents, as technology and product requirements change, leading to variations in the principal's efficient trading partner. A relational contract then specifies how the investment by the principal and the price paid to the agent depend on current costs and the history of the game.

In Section 3, we suppose that the principal commits to a specified investment strategy, allowing us to focus on the agents' incentives. When the principal invests in agent i, the agent has the option to hold up the principal. In order to prevent this opportunism, the principal must give the agent a sufficiently large rent, so that he resists the temptation. This rent can come both in terms of payment today and promised payments in future periods. Crucially, the principal can use these delayed payments in order to prevent future incidences of holdup. This means that the principal must only pay the agent one rent, independent of the number of times they trade. That is, the rent acts like a fixed cost of initiating a new relationship.

When designing the contract, the principal must tradeoff exploiting the gains from trade against the cost of starting new relationships. We show that the principal's profit–maximising contract therefore divides agents into "insiders", with whom she has previously traded, and "outsiders", with whom she has never traded. The principal then uses insiders efficiently, while being biased against outsiders. The model thus provides a theory of endogenous switching costs, where there is a cost to use a new agent, but no cost to return to an old agent.

A comparison with efficiency wage models is informative. Shapiro and Stiglitz (1984) showed that an agent can be persuaded not to hold up a principal if they are awarded rents. Greif (1993) subsequently observed that an agent who trades more frequently demands lower per–period payments. Greif's result is driven by the effect we identified above: by spreading payments over time, the rents used to prevent an agent cheating today can be also used to prevent him cheating tomorrow. The current paper uses this observation to explicitly identify the cost associated with a new relationship. Moreover, in contrast to the efficiency wage model, we solve for the principal's optimal hiring policy, rather than forcing her to pick randomly from the pool of unemployed agents.

The insider–outsider contract has one large defect: it is not time consistent. From the principal's time–0 perspective, each agent must be paid a rent for the first trade, but not for

[2]The principal is female, while the agents are male.

subsequent trades. This has the consequence that insiders trade efficiently. The problem is that many of the agents' rents come from promised future payments, designed to prevent holdup during later trades. The principal thus has an incentive avoid trade with insiders in later periods in order to avoid these promised future payments. Anticipating the principal will renege in later periods, the agent demands more compensation in earlier periods. This may cause the principal to renege even more frequently, leading the agent to increase his demands once again.

The time inconsistency has two implications. First, when the principal is impatient, she will be tempted to default. In Section 4, we examine this problem, supposing that the relationship is enforced by bilateral punishment: that is, a defection by the principal against agent i is only observed by agent i. We show that when the principal is sufficiently patient and costs are IID, the insider–outsider contract is self–enforcing. This is even the case if there are an infinite number of agents. Intuitively, each agent must be paid a rent, so the principal will only ever wish to trade with a finite number of partners. As the discount factor approaches 1, the maximum number of insiders grows to infinity, but it grows sufficiently slowly that the profit per agent increases without bound.

The second implication of time inconsistency is that the principal has an incentive to exaggerate the cost of investing in insiders in order to avoid trading with them as frequently. In Section 5, we suppose that the principal's costs are private information and show that the principal can overcome this problem by using an 'employment contract', where insiders are paid a fixed amount per period, independent of whether trade occurs or not. The advantage of the employment contract is that, since the payments to insiders are sunk, the principal has no incentive to misstate her true costs. The disadvantage is that the principal has more incentive to renege than when using the optimal prices from Section 4. Nevertheless, under IID costs, the employment contract is self–enforcing when the principal is sufficiently patient.

Our results rely on the fact that agents can guarantee themselves rents by threatening to hold up the principal. However, if the principle can demand up–front payments then she can extract all rents from the agents. In Section 6 we argue that such up–front payments are not feasible if the agents are poor, or if there is free entry into the market for principals.

This paper builds on a number of relational contracting papers. MacLeod and Malcomson (1989) examine a repeated moral hazard problem with one worker and deterministic output. Levin (2003) extends the model to allow for either moral hazard over the worker's effort or adverse selection over the worker's type. Levin (2002) allows the principal to employ multiple workers, comparing multilateral and bilateral contracts. These papers differ from ours in that they all assume that contractible transfers are feasible. As a result, the profit–maximising contract is stationary and maximises joint surplus.

Thomas and Worrall (1994) analyse a holdup game where agents receive rents. They allow for different levels of investment, but assume there is only one agent. Thomas and Worrall show that the optimal contract involves a gradual increase of investment, eventually converging to the efficient level. The latter result is related to our finding that insiders are treated efficiently.[3]

In independent work, Calzolari and Spagnolo (2006) also analyse a firm who procures from multiple suppliers under incomplete contracts. Calzolari and Spagnolo assume the principal uses a stationary contract, choosing the cheapest of $n < N$ suppliers, and that prices are determined via an auction. In contrast, we solve for the principal's optimal contract, determining prices and allocations jointly.

Our paper also builds on community enforcement games, as pioneered by Kandori (1992), whereby a population of agents are randomly matched into pairs, and then play a partnership game. In this vein, Sobel (2006) supposes that relationships may grow stale, and agents then have the choice to continue or quit. As in the current paper, Sobel shows that relationships last longer than under complete contracting. However, the paper does not solve for the optimal contract, and does not allow a player to sustain multiple relationships simultaneously.[4]

The paper is organised as follows. Section 3 derives the principal's profit–maximising contract when she can commit to a contingent investment plan. Section 4 examines whether this contract is self–enforcing when the principal cannot commit. Section 5 extends the model to allow the principal to privately observe her cost structure. Section 6 discusses the feasibility of rent–extraction contracts, while Section 7 concludes.

2 The Environment

The economy consists of a principal and N agents. Time $t \in \{0, 1, 2, \ldots\}$ is discrete and infinite. At time 0, the principal designs a contract to maximise her expected profits. Each period $t \in \{1, 2, \ldots, \}$ then consists of three stages (see Figure 1).

1. The cost of investing in agent i, $c_{i,t} \in [\underline{c}, \overline{c}] \subset [0, \infty)$, is publicly revealed. These costs are distributed according to some stochastic process.

2. The principal chooses to invest in at most one of the agents. Denote the probability that the principal invests in i by $Q_{i,t} \in \{0, 1\}$, where $\sum_i Q_{i,t} \leq 1$.

3. The winning agent produces and sells a product worth v. He then chooses to keep $p_t \in [0, v]$ and pay the principal $v - p_t$.

We assume that prices p_t are observable but not contractible, either because courts are inefficient or because the transfer is unverifiable. The stage game of this model thus exhibits

[3] Also see Albuquerque and Hopenhayn (2004), and DeMarzo and Fishman (2007).

[4] On a related note, Ramey and Watson (1997) and Sigouin (2003) show that the desire to maintain a relationship can increase long–term specific investment above the first–best level, easing the dynamic enforcement constraints.

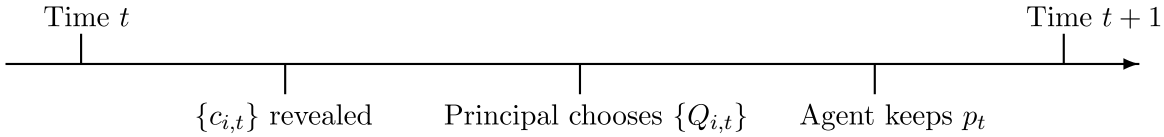

Figure 1: **Timeline.**

the holdup problem. The agent has all the ex–post bargaining power, so will expropriate all the quasi–rents, v. Anticipating being held up, the principal will then abstain from investing.

This paper considers the infinitely repeated version of the holdup game, where all parties have discount rate $\delta \in (0,1)$. We aim to model a decentralised market, so assume that contracts are maintained by bilateral punishment: that is, a deviation in the relationship between the principal and agent i cannot be observed by agents $j \neq i$.[5] Formally, at time t, agent i observes $h_{i,t} := \{c_{1,t}, \dots, c_{N,t}, Q_{i,t}, p_t Q_{i,t}\}$. The agent's history at time t is thus $h_i^t := (h_{i,1}, \dots, h_{i,t})$. Similarly, the principal's history is $h^t := (h_1^t, \dots, h_N^t)$. The principal's investment strategy is a mapping $Q_{i,t} : h_i^{t-1} \times [\underline{c}, \overline{c}]^N \to \{0,1\}$, while the winning agent's pricing strategy is a mapping $p_t : h_i^{t-1} \times [\underline{c}, \overline{c}]^N \to [0, v]$.

A relational contract $\langle Q_{i,t}, p_t \rangle$ is defined to be a history–contingent plan of allocations and prices. In Section 3 we assume the principal commits to her investment strategy, allowing us to focus on the agents' incentives. A contract is said to be *agent–self–enforcing* (ASE) if the agents' strategies form a subgame perfect equilibrium, taking the principal's strategy as given. In Section 4 we analyse the full problem, where the principal cannot commit to her strategy. A contract is then said to be *self–enforcing* (SE) if both the agents' and the principal's strategies form a subgame perfect equilibrium. Among the class of self–enforcing contracts, we then look for the contract that maximises the principal's profit.

Welfare at time t from the principal's relationship with agent i is

$$W_{i,t} := E_t \left[\sum_{s=t}^{\infty} \delta^{s-t} (v - c_{i,s}) Q_{i,s} \right]$$

where "E_t" is the expectation in period t, after $\{c_{i,t}\}$ have been revealed. Total welfare is $W_t := \sum_i W_{i,t}$. Agent i's utility at time t is

$$U_{i,t} := E_t \left[\sum_{s=t}^{\infty} \delta^{s-t} p_s Q_{i,s} \right] \tag{2.1}$$

[5] Allowing for multilateral punishments does not change the optimal contract in Section 3, and simplifies the analysis of Section 4.

The principal's profit at time t from agent i is

$$\Pi_{i,t} := E_t \left[\sum_{s=t}^{\infty} \delta^{s-t} (v - c_{i,s} - p_s) Q_{i,s} \right] \tag{2.2}$$

The principal's total profit is $\Pi_t := \sum_i \Pi_{i,t}$.

2.1 Discussion of the Model

For simplicity, we have phrased the model in terms of the agent stealing from the principal. This stage game can be interpreted far more broadly.

First, consider a moral hazard interpretation. For example, suppose the government wishes to hire a computer contractor to complete a project. The timing is as follows: First, the principal chooses an agent; Second, the principal invests in the equipment the agent advises it to buy at cost c_i; Third, the agent chooses whether to exert non–contractible effort, $e = v - p$. In the computer procurement example, this effort may take the form of providing well qualified staff and coming up with innovative solutions, as suggested by Kelman (1990).

Second, consider a Grossman–Hart–Moore interpretation. For example, suppose Toyota wishes to contract with a supplier for a certain component. The timing is as follows: First, the principal chooses an agent with whom to trade; Second, the principal makes relational–specific investment at cost c_i; Third, the principal and agent decide whether or not to trade, creating value v. At this final stage, assume the agent has all the bargaining power and the parties cannot commit not to renegotiate. In this environment, the parties could theoretically write a contract about the price prior to investment. However, as argued by Hart (1995, Chapter 4.1), such a contract would be useless: the agent could always get out of the contract and renegotiate to a price $p = v$.[6]

3 One–Sided Commitment

In this section we suppose that the principal can commit to her investment strategy, $Q_{i,t}$. We then solve for the optimal agent–self–enforcing contract.

[6]The argument is as follows. Suppose the parties wish to trade one of several widgets. The identity of the useful widget is not known ex–ante and is not observed by the court; the other widgets are useless. First, consider a contract where the principal chooses the widget, and the agent can accept or reject. If $p < v$, the agent will reject and renegotiate, obtaining $p = v$. Second, consider a contract where the agent chooses the widget, and the principal can accept or reject. In this case, the agent will choose a useless widget, which will be rejected by the principal. The parties will then renegotiate and set $p = v$. See Hart and Moore (1999) and Segal (1999) for a broader analysis.

Lemma 1. *A contract $\langle Q_{i,t}, p_t \rangle$ is agent–self–enforcing if and only if*

$$(U_{i,t} - v)Q_{i,t} \geq 0 \quad (\forall i)(\forall t) \tag{DEA}$$

Proof. Fix $\langle Q_{i,t}, p_t \rangle$ and assume that (DEA) holds. Suppose the principal uses the grim trigger strategy: if i deviates at time t, then $Q_{i,s} = 0$ for $s > t$. If $Q_{i,t} = 1$, the agent will not deviate since $U_{i,t} \geq v$. If $Q_{i,t} = 0$, the agent has no action and cannot deviate. The contract is thus agent–self–enforcing.

Next, suppose that (DEA) does not hold. There exists a time t, when $Q_{i,t} = 1$ and $U_{i,t} < v$. The agent will therefore defect and obtain v. □

The agent's dynamic enforcement constraint (DEA) is necessary and sufficient for the agent not to defect. Sufficiency follows from the use of the grim–trigger strategy. Necessity, follows from the fact that the grim–trigger punishment attains the agent's min–max, so is an optimal penal code (Abreu (1988)).

The *principal's problem* is to choose the contract $\langle Q_{i,t}, p_t \rangle$ to maximise time–0 profit

$$\Pi_0 = E_0 \left[\sum_{s=1}^{\infty} \sum_i \delta^s (v - c_{i,s} - p_s) Q_{i,s} \right] \tag{3.1}$$

subject to the agents' dynamic enforcement constraint (DEA). This problem can be simplified by noting that the principal's profit equals welfare minus the sum of agents' utilities. The principal's problem is then to maximise

$$\Pi_0 = E_0 \left[\sum_{s=1}^{\infty} \sum_i \delta^s (v - c_{i,s}) Q_{i,s} \right] - \sum_i U_{i,0} \tag{3.2}$$

subject to (DEA).

Let $\tau_i(t) := \min\{s \geq t : Q_{i,s} = 1\}$ be the time agent i next trades. Lemma 2 then provides a lower bound on agents' utilities at each time period.

Lemma 2. *A contract $\langle Q_{i,t}, p_t \rangle$ is agent–self–enforcing if and only if*

$$U_{i,t} \geq E_t[v\delta^{\tau_i(t)-t}] \quad (\forall i)(\forall t) \tag{DEA$'$}$$

Proof. (DEA$'$) follows from (DEA) and $U_{i,t} = E_t[U_{i,\tau_i(t)} \delta^{\tau_i(t)-t}]$. □

The principal's problem is thus to choose $\langle Q_{i,t}, p_t \rangle$ to maximise profit (3.2) subject to (DEA$'$). Agents' utilities only enter profit via the time–0 term so, if it is feasible, profit is maximised by ensuring that (DEA′) binds at time–0,

$$U_{i,0} = E_0[v\delta^{\tau_i(0)}] \tag{3.3}$$

In Section 3.2 we show that this is indeed feasible: there exist prices $p_t \in [0, v]$ such that (3.3) holds, and (DEA′) is satisfied at all other times. The principal's problem therefore reduces to choosing $\langle Q_{i,t} \rangle$ to maximise profit

$$\Pi_0 = E_0\left[\sum_{s=1}^{\infty}\sum_i \delta^s(v - c_{i,s})Q_{i,s}\right] - E_0\left[\sum_i v\delta^{\tau_i(0)}\right] \tag{3.4}$$

We thus solve the problem in two steps. First, we choose the allocations $Q_{i,t}$ to maximise profits (3.4). Second, we use the agents' dynamic enforcement constraints to derive prices.

3.1 The General Solution

The principal's problem (3.4) is an optimal–stopping problem and is hard to fully characterise. Fortunately, a number of economically relevant results are straightforward to derive.

Denote the set of *insiders* at time t by $\mathcal{I}_t := \{i : \tau_i(0) < t\}$.

Proposition 1. *An optimal agent–self–enforcing contract has the following properties:*
(a) Trade with insiders is efficient. Suppose $i \in \mathcal{I}_t$. Then $Q_{i,t} = 1$ if

$$c_{i,t} < v \quad \textit{and} \quad c_{i,t} < c_{j,t}\ (\forall j) \tag{3.5}$$

(b) Trade is biased against outsiders. Suppose $i \notin \mathcal{I}_t$. Then $Q_{i,t} = 0$ if either

$$(i) \qquad (v - c_{i,t}) < v(1-\delta); \quad \textit{or} \tag{3.6a}$$

$$(ii) \qquad (c_{j,t} - c_{i,t}) < v(1-\delta) \quad \textit{for some } j \in \mathcal{I}_t \tag{3.6b}$$

Proof. See Appendix A.1 □

Part (a) says that, if an insider has a lower cost than all other agents, then he will be chosen by the principal. In addition, the principal treats all insiders equally: conditional on trading with an insider, she always chooses the cheapest.[7] The intuition for these results is as follows. Each agent obtains rents of v the first time they trade, by threatening to hold up the principal. Agents care about the lifetime value of their utility so, by delaying payments to the agent, the principal can use the same sum of money to avoid hold up during later trades with the same agent. This means that the agent only earns one rent, no matter how many times he trades.

[7] This follows from equation (A.1).

This rent acts like a fixed cost of initiating a new relationship; once the fixed cost has been paid, the principal uses the agent efficiently.

Part (b) says that trade is biased against outsiders. There are three types of biases. First, the principal may abstain rather than trade with an outsider, $i \notin \mathcal{I}_t$. If the principal refrains from investing in i, she loses the gains from trade, $v - c_{i,t}$. However, she postpones paying the agent's rent by at least one period, gaining the rental value of the rent, $(1-\delta)v$. This yields equation (3.6a).

The second bias is that the principal may prefer an insider, j, over an outsider, i. If the principal trades with the insider rather than the outsider, she loses the cost difference, $c_{j,t} - c_{i,t}$. However, she also delays the rent by at least one period, gaining $(1-\delta)v$. This yields equation (3.6b). This result means that an agent who enters the market late will struggle to trade with the principal. For example, Proposition 4 show that, when costs are IID, the insider advantage increases monotonically with the number of insiders, and eventually grows so large than no more outsiders are used.

The third bias is that the principal may choose to trade with a relatively inefficient outsider. For example, there is a bias towards agents who are more efficient on average. This is formalised by the following result.

Proposition 2. *Assume $N < \infty$ and each agent $i \in \{1, \dots, N\}$ has costs $c_{i,t}$ that are IID draws from some i–specific distribution. Suppose that the distribution of $c_{1,t}$ is smaller than $c_{2,t}$ in terms of first–order stochastic dominance. If, in any state of the world, $c_{1,t} = c_{2,t}$ then the principal prefers to use agent 1 over agent 2.*

Proof. See Appendix A.2 □

The intuition behind Proposition 2 is that the principal cares both about the cost in the current period and in future periods, after the agent has become an insider, so she prefers an agent who is more efficient on average, even if they are not the most efficient in the current period. Empirically, this suggests a bias towards larger, more established firms. For example, Kelman (1990, p. 8) finds the largest computer vendor, IBM, had more than twice the market share among private companies than with the government. The result also suggests a bias towards generalists rather than specialists. Indeed, many households use a general contractor rather than a series of specialists (e.g. plumbers, painters, carpenters) in order to reduce their exposure to holdup.

Proposition 2 has the corollary that two agents may like to merge in order to make themselves larger and give themselves a competitive advantage. To illustrate, suppose costs are IID and that agents 1 and 2 can merge, causing them to act as a single entity and giving them joint cost of $\min\{c_1, c_2\}$. Then, if N is sufficiently large, such a merger will increase their joint utility.[8]

[8] Proof. Before 1 and 2 are merge, they have $U_{i,0} = O(1/N)$ as $N \to \infty$. Next, suppose agents 1 and

When choosing among outsiders, the principal also prefers more stable suppliers, who are expected to be in the market for a long time, over short–term players. Intuitively, when using a stable supplier, the principal can amortise the rents over a greater number of periods. This means that longevity can become self–fulfilling: firms who are expected to leave the market obtain no business and have no incentive to stay.

The optimal contract in Proposition 1 resembles the supply associations in Japanese industry. For example, Toyota invites preferred suppliers (defined by size, dependency and performance) to join its supply associations (e.g. Kyohokai, Seihokai, Eihokai). These suppliers account for a large fraction of Toyota's parts supply (typically over 80%) and are implicitly promised a share of future business. As in Proposition 1, new firms occasionally join, but old firms rarely leave. In 1992, Toyota's Tokai Kyohokai consisted of 141 members; over the previous 20 years, 29 firms joined but only 5 left (**?**,Asanuma (1989), Sako (1996)).

A few general points about Proposition 1 are worth emphasising. First, it provides a foundation for switching costs. However, unlike typical models of switching costs (e.g. Klemperer (1995)), there is a cost to using a new partner for the first time, but not to returning at a later date. Second, the optimal contract is not stationary, as assumed in efficiency wage models (e.g. Shapiro and Stiglitz (1984)). Third, we made no assumptions about the stochastic process governing costs. With more structure, one can strengthen the conditions in part (b).

3.2 Prices

The profit–maximising investment function $Q_{i,t}$ is chosen to maximise profits (3.4). Prices, $p_t \in [0, v]$, can then be chosen in a number of ways such that (a) utilities satisfy the agents' dynamic enforcement constraints (DEA) and, (b) the principal gives no rents away (3.3). Given agents' utilities, prices are then determined by backwards induction:

$$p_t = U_{i,t} - E_t[\delta^{\tau_i(t+1)} U_{i,\tau_i(t+1)}] \tag{3.7}$$

whenever $Q_{i,t} = 1$.

Adopting the terminology of Hart and Moore (1994), the *fastest prices* have the property that the dynamic enforcement constraints (DEA) bind in all periods. These prices are so called because the principal makes payments to the agent as early as possible. Using (3.7) the fastest prices are given by

$$p_t = v\, E_t[1 - \delta^{\tau_i(t+1)}] \tag{3.8}$$

2 merge. Suppose the principal has no insiders. If she uses firm $i \notin \{1, 2\}$, her profit is at most $-\underline{c} + \Pi_1$, where Π_n is the profit from n insiders. If she uses the new merged firm, her profit is $-\min\{c_1, c_2\} + \Pi_2$. Observe that an extra insider is valuable in the states when it is unprofitable to bring in an outsider: that is, $\Pi_2 - \Pi_1 \geq E[c_1 - \min(c_1, c_2)] \Pr(c_1 - \underline{c} < v(1 - \delta)) =: k > 0$. Hence the principal chooses the merged firm whenever $\min\{c_1, c_2\} - \underline{c} < k$. As a result, the agents' joint utility is bounded above zero as $N \to \infty$.

whenever $Q_{i,t} = 1$. With these prices, the time of the next trade is a sufficient statistic for the price. In addition, the agent bears much of the risk in the contract since the price depends on the expected time until the next trade, yet is only paid if that trade occurs.

In contrast, the *slowest prices* have the property that the principal delays payment by as much as possible. This means that there exists a t_i^* such that, whenever $Q_{i,t} = 1$,

$$\begin{aligned} p_t &= 0 && \text{for } t < t_i^* \\ &\in (0, v] && \text{for } t = t_i^* \\ &= v && \text{for } t > t_i^*. \end{aligned}$$

The problem with the slowest prices is that, if the principal cannot commit to her investment strategy, then she will be tempted to defect at time t_i^*. This following result formalises this idea.

Proposition 3. *Fix an allocation rule $Q_{i,t}$. Denote the profit under the fastest prices by $\Pi_{i,t}$ and the profit under any other price system by $\Pi'_{i,t}$. Then $\Pi_{i,t} \geq \Pi'_{i,t}$ $(\forall i)(\forall t)$.*

Proof. Profit at time t is given by $\Pi_{i,t} = W_{i,t} - U_{i,t}$. Utilities must obey (DEA′), so $\Pi_{i,t}$ is maximised by setting $U_{i,t} = E_t[v\delta^{\tau_i(t)-t}]$. □

Proposition 3 shows that the fastest prices maximise the principal's profit in each period. Consequently, when the principal cannot commit to the contract (Section 4), an investment rule is implementable only if it can be implemented by the fastest prices.

3.3 IID Costs

In this section we suppose costs are identically and independently distributed both over time and across agents, with support $[\underline{c}, \overline{c}]$. The IID model is appealing because the identity of most efficient agent constantly changes, emphasising the tradeoff between exploiting the gains from trade and maintaining a long run relationship. Since the number of insiders can be treated as the state variable, the IID case is also very tractable.

First, we can characterise the maximum number of insiders, n^*. Suppose there are $n^* - 1$ insiders, and costs are most favourable for using an outside agent. Then it must be the case that the principal prefers to use one last outsider and pay rent v, rather than sticking to the current set of insiders. That is,

$$-\underline{c} + \frac{\delta}{1-\delta}\omega(n^*) \geq \max\{v - \overline{c}, 0\} + \frac{\delta}{1-\delta}\omega(n^* - 1) \tag{3.9}$$

where $\omega(n) := E[\max\{v - c_{1:n}, 0\}]$ is the expected welfare from n agents and $c_{1:n} := \min\{c_1, \ldots, c_n\}$. Similarly, it must be the case that the principal does not want to increase the number of agents from n^* to n^*+1, yielding another equality analogous to (3.9). Together these inequalities imply that n^* is the unique integer satisfying

$$\frac{\delta}{1-\delta}\Delta\omega(n^*) \geq \max\{v-\overline{c},0\}+\underline{c} \geq \frac{\delta}{1-\delta}\Delta\omega(n^*+1) \tag{3.10}$$

where $\Delta\omega(n) := \omega(n)-\omega(n-1)$.[9] Since the number of insiders increases over time, $n_t \to n^*$ almost surely. Moreover, as agents become more patient, the benefit from future cost reductions increases and n^* rises.

Next, we can characterise the principal's investment strategy. At any time t, we can summarise the state of the world by the number of insiders, $n_t := |\mathcal{I}_t|$. The evolution of n_t is then described by a time–invariant markov chain. The transition depends upon the lowest costs of the insiders and outsiders, denoted c^{I} and c^{O} respectively. Since the principal can always choose not to invest, the effective inside cost is $\tilde{c}^{\mathrm{I}} := \min\{v, c^{\mathrm{I}}\}$.

The transition process can be characterised by backwards induction, using n^* as an initial condition. Abusing notation, denote the value function of the principal's profit–maximisation problem in state n by Π_n. This evolves as follows:

$$\begin{aligned} \Pi_n &= E[v-\tilde{c}^{\mathrm{I}}]+\delta\Pi_n && \text{if } i\in\mathcal{I}_t\cup\emptyset \\ &= E[-c^{\mathrm{O}}]+\delta\Pi_{n+1} && \text{if } i\in N\backslash\mathcal{I}_t \end{aligned} \tag{3.11}$$

Denote the insider cost differential by $\Delta := \tilde{c}^{\mathrm{I}}-c^{\mathrm{O}}$. The optimal policy opts for the inside agent if the cost differential is less than some cutoff, $\Delta < \Delta_n^*$. The value function then becomes

$$\Pi_n = E[-c^{\mathrm{O}}]+E[\delta\Pi_{n+1}\mathbf{1}_{\Delta>\Delta_n^*}]+E[(v-\Delta+\delta\Pi_n)\mathbf{1}_{\Delta\leq\Delta_n^*}] \tag{3.12}$$

Using (3.11), the cutoff Δ_n^* is chosen so that

$$(v-\Delta_n^*) = \delta\left(\Pi_{n+1}-\Pi_n\right) \tag{3.13}$$

Using (3.12) to substitute for Π_n yields,

$$(v-\Delta_n^*) = \frac{\delta}{1-\delta}\left[(1-\delta)\Pi_{n+1}-E\left[\max\{v-\Delta_n^*, v-\Delta\}-c^{\mathrm{O}}\right]\right] \tag{3.14}$$

The left hand side of equation (3.14) equals today's lost profits from investing in an extra agent. The right hand side equals the future discounted gains from using an extra agent, taking into account the fact a principal who uses an insider this period retains the right to use an outsider

[9] Proof that $\Delta\omega(n)$ decreases in n. Let $x_0 = \max\{v-c_{1:n-1},0\}$, $x_1 = v-c_n$ and $x_2 = v-c_{n+1}$. Then $\Delta\omega(n+1) = E[\max\{x_2-\max\{x_0,x_1\},0\}] \geq E[\max\{x_2-x_0,0\}] = E[\max\{x_1-x_0,0\}] = \Delta\omega(n)$, since x_1 and x_2 have the same distribution.

in the future, if $\Delta > \Delta_n^*$.

We can now derive the optimal policy for any distribution of costs and any number of agents. First, solve for the maximal number of agents, n^*, and calculate the associated value function. Second, use equation (3.14) to solve for the optimal state–$(n^* - 1)$ cutoff and equation (3.12) to calculate the value function Π_{n^*-1}. Then iterate.

Table 1 provides a numerical illustration. In this example there are an infinite number of agents willing to trade with the principal. Despite all this choice, the principal only ever uses six agents. The principal's adoption of new agents rapidly diminishes over time: it takes an average of 4 periods for her to use 3 different agents, and 35 periods to use 5 different agents. This distorted trading pattern lowers the principals profits to \$83.6, relative to \$100 under the first–best.

Insiders, n_t	Cutoff, Δ_n^*	Prob$(n_{t+1} = n_t)$	Value function, Π_n
0	0	0	83.6
1	0.358	0.358	85.3
2	0.398	0.637	87.0
3	0.454	0.837	88.6
4	0.549	0.959	90.2
5	0.834	0.999	91.7
6	1	1	92.9

Table 1: **The Evolution of Trade.** This example assumes the value of trade is $v = 2$, costs are $c_{i,t} \sim [0, 1]$, the number of suppliers is $N = \infty$, and the discount rate is $\delta = 0.98$.

More generally, we can show the following comparative static.

Proposition 4. *Suppose agents' costs are IID, and that either $\underline{c} > 0$ or $v > \overline{c}$. Then Δ_n^* and Prob$(n_{t+1} = n_t | n_t = n)$ increase in n.*

Proof. In Appendix A.3 we show Δ_n^* increases in n. Since the number of insiders increases in n, the number of outsiders decreases in n and Δ_n^* increases in n, we conclude that Prob$(n_{t+1} = n_t | n_t = n)$ increases in n. □

Proposition 4 shows that, as the number of insiders grows, so the gain from an extra insider declines and the optimal cutoff increases.

Throughout this section we have assumed that the principal can commit to her strategy. This assumption is not unreasonable: Example 2 in Section 4.1 shows that the principal will not renege on her investment plan under the assumptions in Table 1. Next, we investigate this issue in more detail.

4 No Commitment

In this section we examine the optimal self-enforcing contract. To motivate the discussion, example 1 shows that the principal's dynamic enforcement constraint may strongly restrict the set of possible trading opportunities.

Example 1 (Unravelling Trade). Suppose $N = 1$ and $c_{1,t} \in (0, v)$ are deterministic and increasing such that $\lim_{t\to\infty} c_{1,t} > \delta v$ and $\sum_{t=1}^{\infty} \delta^t (v - c_{1,t}) > v$. See Figure 2 for an illustration. In this case, the optimal ASE contract obeys $Q_{1,t} = 1$ $(\forall t)$. However, since gains from trade disappear over time, any contract eventually satisfies $W_{1,t} < v$ for $t \geq t^*$. Thus trade is not feasible at time t^* and, by backwards induction, the only solution to satisfy (**??**) is $Q_{1,t} = 0$ $(\forall t)$. △

As Example 1 shows, the insider-outsider contract is not time consistent. From the principal's time-0 perspective, she has to pay rents v to an agent in order to stop him defecting. After these rents have been paid, the principal uses the agent efficiently, as shown in Proposition 1. The problem is that many of these rents come in terms of promised future utility which are needed to stop future defections. These postponed payments mean that the principal may later regret promising to use the agent efficiently. If the principal defects in these later periods, she then raises the required price in earlier periods, exerting a negative externality on her former self.

The relational contract turns around the holdup problem. In a one-shot game, the principal is concerned that the agent will cheat after the specific investment. This problem is solved through a long term relationship, by delaying the payment of the agent's rents. This relational contract then introduces a second holdup problem: the principal is tempted to defect, and refuse to pay these promised rents.

The problem of time consistency is very real for car manufacturers. Throughout the 1980s, GM ostensibly adopted the Japanese partnering model. However, at the turn of the millennium, GM was able to source from China, and "jumped to the conclusion that the immediate benefits of low wage costs outweighed the long-term benefits of investing in relationships" (Liker and Choi (2004)).

4.1 IID Costs with Patient Agents

When the principal cannot commit to the contract, a contract $\langle Q_{i,t}, p_t \rangle$ is self–enforcing if and only if it satisfies the agents' dynamic enforcement constraint (DEA) and the principal's own dynamic enforcement constraint (DEP).

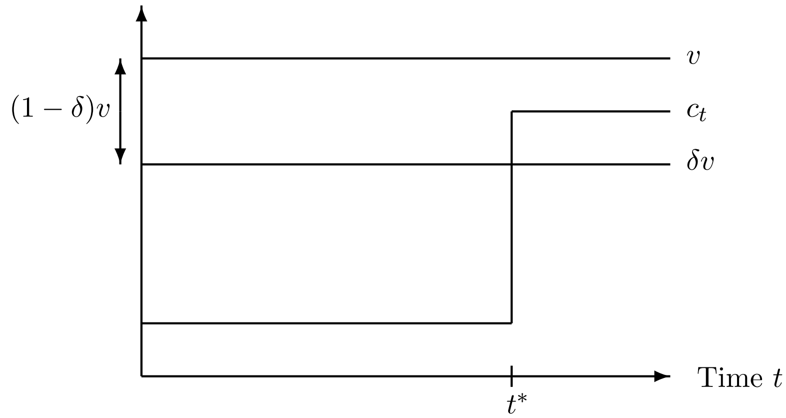

Figure 2: **Time Inconsistency of the Optimal ASE Contract.** In this example, the optimal ASE stipulates trade in every period. However, the increase in costs at time t^* causes the principal to quit the relationship, leading the contract to unravel.

When relationships are enforced by bilateral punishments, a defection by the principal on agent i will be punished by the termination of relationship i. Hence the constraint is

$$\Pi_{i,t} Q_{i,t} \geq 0 \quad (\forall i)(\forall t). \tag{DEP$_\mathrm{B}$}$$

Appealing to Proposition 3, (DEA$'$) binds in each period, enabling us to rewrite (DEP$_\mathrm{B}$) as

$$\Pi_{i,t} = (W_{i,t} - v) Q_{i,t} \geq 0 \quad (\forall i)(\forall t) \tag{DEP$'_\mathrm{B}$}$$

Bilateral enforcement is realistic in large decentralised market where agent i's information about agent j's relationship is poor. For example, in Vietnam, McMillan and Woodruff (1999, p. 648) report that only 19% of interviewees thought that a cheating customer would be blacklisted by other firms in the industry. Similarly, in the U.S., firms often refuse to report negative information about former suppliers or employees for fear of legal action (Kelman (1990, p. 49), Bewley (1999)).

When relationships are enforced by multilateral punishments, a defection by the principal on agent i will be punished by the termination of all relationships. Hence the constraint is

$$\Pi_t \sum_i Q_{i,t} \geq 0 \quad (\forall i)(\forall t). \tag{DEP$_\mathrm{M}$}$$

Appealing to Proposition 3, (DEA$'$) binds in each period, enabling us to rewrite (DEP$_\mathrm{M}$) as

$$\Pi_t = E_t \left[\sum_{s=t}^{\infty} \sum_i \delta^{s-t} (v - c_{i,s}) Q_{i,s} \right] - E_t \left[\sum_i v \delta^{\tau_i(t)} \right] \geq 0 \quad (\forall i)(\forall t). \tag{DEP$'_\mathrm{M}$}$$

The multilateral constraint is weaker than the bilateral constraints, and therefore increases the profits of the principal. Indeed, **?** claims that Toyota's supply associations were formed to share information and safeguard against opportunism.

As in Section 3, the *principal's problem* reduces to choosing $\langle Q_{i,t}\rangle$ to maximise time-0 profit (3.1) subject to ($\mathrm{DEP}'_{\mathrm{B}}$) or ($\mathrm{DEP}'_{\mathrm{M}}$). Given the optimal investment rule, payments are given by the fastest prices (3.8). Proposition 5 shows that, when the principal is sufficiently patient, the optimal contract is self–enforcing.

Proposition 5. *Suppose that costs are IID, and that either $\underline{c} > 0$ or $v > \bar{c}$. Then there exists $\hat{\hat{\delta}} < 1$, independent of N, such that the optimal agent-self-enforcing contract satisfies ($\mathrm{DEP}_{\mathrm{B}}$) and ($\mathrm{DEP}_{\mathrm{M}}$) when $\delta > \hat{\hat{\delta}}$.*

Proof. See Appendix A.4. □

Example 2 (Uniform Costs). Suppose $N = \infty$, $c_{i,t} \sim U[0,1]$ and $v > 1$. As shown is Appendix A.5, ($\mathrm{DEP}_{\mathrm{B}}$) and ($\mathrm{DEP}_{\mathrm{M}}$) are satisfied if $\delta \geq \hat{\hat{\delta}} := 1/(1 + (v-1)^3)$. For example, if $v = 2$, then $\hat{\hat{\delta}} = 1/2$, so the example in Table 1 is self–enforcing. As $v \to 1$, so $\hat{\hat{\delta}} \to 1$. △

The bilateral dynamic enforcement constraint requires that, when the principal trades with agent i, the future profits associated with i are positive. For fixed N, this trivially holds as the discount rate approaches one and profit per agent tends to infinity. Hence there exists a $\hat{\hat{\delta}}_N$ such that the optimal contract satisfies ($\mathrm{DEP}_{\mathrm{B}}$) when $\delta \geq \hat{\hat{\delta}}_N$.

Proposition 5 makes a stronger statement: there exists a critical discount factor, $\hat{\hat{\delta}}$, independent of the number of agents, such that the optimal contract satisfies ($\mathrm{DEP}_{\mathrm{B}}$) when $\delta \geq \hat{\hat{\delta}}$. With an infinite number of agents, the problem is that the number of trading partners will increase without bound as $\delta \to 1$. Despite this, the average profit per agent tends to infinity as the parties become more patient. Intuitively, the principal's investment in new partners is limited by having to pay rent v for every new relationship. This means that the maximum number of insiders is determined by the marginal benefit of an extra agent, $E[c_{1:n-1} - c_{1:n}]$, as in equation (3.9). This marginal benefit is of order $o(n)$ and therefore decreases more rapidly than the average benefit of each relationship, $E[v - c_{1:n}]/n$. The average profit per agent thus increases in δ, and the principal will refrain from defecting when sufficiently patient.

4.2 IID Costs with Impatient Agents: Version 1

When the optimal contract (Proposition 4) does not satisfy the principal's dynamic enforcement constraint, she may treat insiders inefficiently in order to increase her continuation profits. The general principal is that the time-0 principal treats the insiders' rents as sunk and therefore ignores them, while the time-t principal still pays these rents, and would therefore like to re-designate the insiders as outsiders.

First, suppose there is one agent, $N = 1$. When there is one insider, $n = 1$, the principal wishes to maximise welfare subject to ($\mathrm{DEP}'_{\mathrm{M}}$). If the constraint binds then the principal can increase her continuation profits by cutting out some low value trades. In particular, trade will occur if and only if $c \leq k_1$, where $k_1 < v$ is defined by the largest root of

$$\frac{\delta}{1-\delta}[(v-c)\mathbf{1}_{c\leq k_1}] - k_1 = 0$$

When $n = 0$, the principal will trade with the agent if and only if $c \leq k_0$. As usual, the optimal contract is biased against outsider: $k_0 < k_1$. Intuitively, a trade at cost k_1 yields zero profit, so the principal will prefer to wait for the cost to fall. In addition, k_0 is less that the corresponding cutoff when the principal can commit.[10] Intuitively, efficient trading opportunities are sometimes forgone after the agent become an insider, reducing the ex-ante value of trade the willingness of the principal to invest in a new agent.

Next, suppose there are two agents, $N = 2$, and the agents use multilateral enforcement. When $n = 2$, the principal wishes to maximise welfare subject to ($\mathrm{DEP}'_{\mathrm{M}}$). If the constraint binds the optimal contract is biased towards insiders who were used more recently. To understand this result, suppose the principal uses agent 1 at time t. At time $t+1$, a small bias towards agent 1 has a second-order effect on welfare but delays the use of agent 2, yielding a first-order reduction of rents. In order to slacken the ($\mathrm{DEP}'_{\mathrm{M}}$) constraint, the principal may also wish to abstain from trade when the insiders costs exceed a cutoff, k_2. When $n = 1$, the principal will be biased against the outsider. Trade with the insider will also be limited if ($\mathrm{DEP}'_{\mathrm{M}}$) binds.

Finally, suppose there are three agents, $N = 3$. When there are $n = 3$ insiders and ($\mathrm{DEP}'_{\mathrm{M}}$) binds, the principal will bias trade towards insiders with whom she has traded more recently. To understand this result, suppose the principal last used the agents $\{1, 2, 3\}$ at times $\{t_1, t_2, t_3\}$, where $t_1 > t_2 > t_3$. At time $t_1 + 1$, biasing allocation towards 1 over 2 and 3 slackens ($\mathrm{DEP}'_{\mathrm{M}}$) in periods $\{t_2 + 1, \ldots, t_1\}$ by reducing agent 2 and 3's rents. Similarly, biasing the allocation towards 2 over 3, slackens ($\mathrm{DEP}'_{\mathrm{M}}$) in periods $\{t_2 + 1, \ldots, t_2\}$, while having no effect on the (DEP') constraint in periods $\{t_2 + 1, \ldots, t_1\}$. ($\mathrm{DEP}'_{\mathrm{M}}$) constraint in periods $\{t_2 + 1, \ldots, t_1\}$.[11] When there are fewer insiders, the principal is biased against outsiders. As above, trade with the insiders will also be affected if ($\mathrm{DEP}'_{\mathrm{M}}$) binds.

When agents use bilateral enforcement, the results are similar. When the ($\mathrm{DEP}'_{\mathrm{B}}$) constraint binds, the principal will bias trade towards agents she used more recently. For example, suppose $N = 3$, all the agents are insiders and the principal last used the agents $\{1, 2, 3\}$ at times $\{t_1, t_2, t_3\}$, where $t_1 > t_2 > t_3$. At time $t_1 + 1$, a small bias in favour of 1 over 2, and 2 over 3 then leads to second-order losses in total welfare, but first-order increases in the welfare from trades with agent 1. It therefore weakens the ($\mathrm{DEP}'_{\mathrm{B}}$) which states that the welfare from

[10] Recall, with commitment trade occurs if $c_{1,t} < c_0$ for an outsider, and $c_{1,t} < v$ for an insider.

[11] Observe that the ($\mathrm{DEP}'_{\mathrm{M}}$) constraints in periods $\{1, \ldots, t_3\}$ are independent of the policy at time $t_1 + 1$.

4.3 IID Costs with Impatient Agents: Version 2

When the optimal contract (Proposition 1) does not satisfy the principal's dynamic enforcement constraint, she may treat insiders inefficiently in order to increase her continuation profits. The general principal is that the time-0 principal treats the insiders' rents as sunk and therefore ignores them, while the time-t principal still pays these rents, and would therefore like to re-designate the insiders as outsiders.

First, suppose there is one agent, $N = 1$. When there are no insiders $n = 0$, the principal awards the good to the agent if his cost is below come cutoff, k_0. When $n = 1$, the principal wishes to maximise welfare subject to (DEP). If the constraint binds then the principal can increase her continuation profits by cutting out some low value trades. In particular, trade will occur if and only if $c \leq k_1$, where $k_1 < v$ is defined by the largest root of

$$\frac{\delta}{1-\delta}[(v-c)\mathbf{1}_{c\leq k_1}] - k_1 = 0$$

As usual, the optimal contract is biased against outsider: $k_0 < k_1$. Intuitively, a trade at cost k_1 yields zero future expected profits, so the principal prefers to wait for the cost to fall. When compared to the commitment case, the introduction of the (DEP) constraint reduces trade. For an insider, the principal forgoes efficient trades in order to satisfy (DEP). This reduces the ex-ante value of trade and the willingness of the principal to invest in a new agent.

Next, suppose $N \geq 2$ and trade is enforces bilaterally. When $n = 0$, the principal trades with the cheapest agent with cost below a cutoff. When $n = 1$, the principal is biased towards the insider and, if (**??**) binds, will forgo some low value trades in order to reduce the incentive to defect. When n is larger and (**??**) binds, the principal is biased against agents used less recently, and even more biased against an outsider. For example, suppose agents $\{1, 2, 3\}$ constitute the set of insiders, and were last used at times $\{t_1, t_2, t_3\}$, where $t_1 > t_2 > t_3$. At time $t_1 + 1$, a bias in favour of agent 1 over agent 2, and agent 2 over agent 3 increases the welfare of agent 1 at time t, and therefore weakens (**??**). Such biases lead to welfare losses from insiders and therefore reduce the principal's time-0 profit, however these are second order when the biases are small, whereas the increase in the trading agent's welfare are of first-order.

Next, suppose $N \geq 2$ and trade is enforces multilaterally. The results are similar, albeit for slightly different reasons. When $n = 0$ or $n = 1$ the analysis is essentially unchanged. When n is larger, the principal is again biased against agents used less recently, and even more biased against an outsider. Using the above example, a small bias towards agent 1 at time $t_1 + 1$ has a second-order effect on welfare, and therefore time-0 profits, but delays the use of agents 2 and 3, yielding a first-order reduction of rents and slackening (**??**) in periods $\{t_2 + 1, \ldots, t_1\}$. Similarly, a small bias towards 2 over 3, results in a second-order loss in welfare, while delaying the use of agent 3, thereby slackening (**??**) in periods $\{t_3 + 1, \ldots, t_1\}$.

5 Private Cost Information

In this section we suppose that costs $c_{i,t} \in [\underline{c}, \overline{c}]$ are privately observed by the principal. Section 5.1 discusses the problem of incentive compatibility, while Section 5.2 suggests a solution.

5.1 Failure of Incentive Compatibility

In Section 4 we showed that, if agents are sufficiently patient, the insider–outsider contract is self–enforcing. It may not, however, be incentive compatible.

Example 3. Suppose $N = 1$, $v = 1$, costs are $c_{1,t} \sim U[0,2]$ and $\delta = 9/10$. In the optimal insider–outsider contract, the principal trades with an outsider if $c_{1,t} \leq 0.80$, and trades with an insider if $c_{1,t} \leq 1$. This contract is self–enforcing and generates prices $p_t = 0.18$. However, the contract is not incentive compatible. If $c_{1,t} \in [0.82, 1]$, the principal has an incentive to overstate her cost in order to avoid trade with an insider. △

Example 3 shows that the principal may exaggerate her cost to avoid trade with an insider. Similarly, she may misstate her costs in order to trade with an outsider over an insider. These problems result from the time–inconsistency of the insider–outsider contract. Intuitively, since many of an agent's rents take the form of delayed payments, the principal has an incentive to avoid trade in these future periods by pretending that her costs are artificially high.

The principal can try to alter her investment strategy, $Q_{i,t}$, in order to obtain incentive compatibility. For example, the contract could make trade compulsory, restrict the agents with whom the principal trades, or impose stationarity. In Section 5.2 we show there is a better solution: if the principal can expand the space of payments, then she can implement the optimal ASE contract.

5.2 Employment Contract

First, let us expand the set of contracts to allow payments from the principal to the agent in any period after the relationship has formed, i.e. $t \geq \tau_i(0)$. That is, we allow transfers in periods when no trade occurs.[12]

Denote the contract by $\langle Q_{i,t}, p_{i,t} \rangle$, where $p_{i,t}$ is the price paid to agent i in period t by $p_{i,t}$. This contract specifies allocations and payments as a function of the principal's cost reports and the parties past actions. As before, the contract is chosen at time–0 to maximise the principal's profits. We then consider the following game.

1. The principal privately observes her costs $\{c_{i,t}\}$. These costs are distributed according to some stochastic process.

[12]This expansion in the space of payments does not change the set of attainable ASE or SE payoffs when costs are publicly observable: see Section 6.1.

2. The principal makes public cost reports $\{\hat{c}_{i,t}\}$. These reports determine the allocation and payments specified by the contract.

3. The principal invests according to $Q_{i,t} \in \{0,1\}$, where $\sum_i Q_{i,t} \leq 1$.

4. The principal makes payments $p_{i,t} \in \mathbb{R}$ to agent i. The winning agent can refuse and keep v. A losing agent can refuse and keep 0.

A relational contract $\langle Q_{i,t}, p_{i,t} \rangle$ is incentive compatible if the principal cannot improve her payoffs by misreporting at stage 2.

Consider the following *employment contract.* Payments are given by

$$\begin{aligned} p_{i,t} &= (1-\delta)v \qquad && \text{if } i \in \mathcal{I}_t \\ p_{i,t} &= 0 && \text{if } i \notin \mathcal{I}_t \end{aligned} \tag{5.1}$$

The investment strategy $Q_{i,t}$ is then chosen to maximise time–0 profits (3.4).

Proposition 6. *Suppose that costs are privately known by the principal. The employment contract (5.1) is an optimal agent–self–enforcing contract and is incentive compatible. Moreover, it is self–enforcing if*

$$W_{i,t} \geq v \qquad \textit{for all } i \in \mathcal{I}_t. \tag{DEP$^{\text{EC}}$}$$

Proof. See Appendix A.7 □

The employment contract pays each insider $(1-\delta)v$ per period, independent of whether trade occurs or not. That is, each 'employee' submits to the authority of the principal, working when they are told, in exchange for a constant wage. The model thereby provides a theory of the firm, where both firm size and growth are determined endogenously.

The benefit of the employment contract is that the principal's payments to insiders are independent of her cost declarations. As a consequence, the principal's investment strategy is incentive compatible. The contract is also detail–free in that agents do not have to know the distribution of costs, so long as they trust the principal to stick to the contract.

The cost of the employment contract is that the principal's dynamic enforcement constraint (DEP$^{\text{EC}}$) is stricter than under the fastest price contract (**??**). The reason for the difference is straightforward: under the employment contract, an insider receives utility $U_{i,t} = v$ whether or not he trades; under the fastest price contract, an insider receives $U_{i,t} = v$ when he trades and receives strictly less in other periods. Nevertheless, when costs are IID and the parties are patient, Proposition 7 shows that the difference between (**??**) and (DEP$^{\text{EC}}$) is relatively minor.

Proposition 7. *Suppose that costs are IID. In addition, assume that either $\underline{c} > 0$ or $v > \overline{c}$. Then there exists a $\hat{\delta} < 1$, independent of N, such that the optimal agent–self–enforcing contract satisfies (DEP^{EC}) when $\delta > \hat{\delta}$.*

Proof. Same as proof of Proposition 5. □

When costs are IID, the future benefit from the relationship with agent i does not depend on whether they are trading or not this period. Hence the additional dynamic enforcement constraints, needed to guarantee the principal pays each insider each period, are not particularly restrictive for high discount factors. This may not be the case for other cost structures: for example, if costs follow a markov chain then the optimal ASE contract may satisfy (**??**) but not ($\mathrm{DEP}^{\mathrm{EC}}$).

6 Agents' Rents

This paper has argued that the threat of holdup enables the agents to extract rents, and that the principal will allocate investment in order to minimise these rents. The idea that agents earn rents has empirical support. Kelman (1990, p. 84) shows that suppliers' margins on private contracts are higher than those on government contracts. Similarly, Liker and Choi (2004) claim that Toyota is more concerned about suppliers' profitability than GM.

Nevertheless, if agents are allowed to make payments before the time of first trade, then the principal can extract all their rents. For example, the contract can require each agent to make a voluntary payment equal to his lifetime rents at time 0. Since this payment is sunk, it reallocates rents but does not affect the subsequent continuation game. The purpose of this section is twofold. In Section 6.1, we show that if agents are poor, or the principal is untrustworthy, these rent–extraction contracts are not feasible.[13] In Section 6.2, we briefly discuss the form of the optimal relational contracts without rents.[14]

6.1 Sources of Rents

If the set of contracts is sufficiently broad, then the principal can extract all rents from the agents. In this section we provide two reasons why such contracts may not be feasible. First, consider a generalised relational contract $\langle \tilde{Q}_{i,t}, \tilde{p}_{i,t}, \rangle$ consisting of allocations and prices. In period 0, the principal designs the contract and makes payment $\tilde{p}_{i,0}$ to agent i. In each subsequent period, the parties play the following game.

1. Costs $\{c_{i,t}\}$ are publicly revealed.

2. The principal chooses to invest according to $\tilde{Q}_{i,t} \in \{0,1\}$, where $\sum_i \tilde{Q}_{i,t} \leq 1$.

[13] These arguments are not new: see Shapiro (1983), Shapiro and Stiglitz (1984) and Mathewson and Winter (1985). Nevertheless, it is important to establish the results formally.

[14] Throughout this section we assume that all payments are voluntary. Contractible or up–front payments provide another method to extract the agents' rents, by requiring an agent to pay the principal v before the principal invests in him. Such contractible payments do not affect the results in this section.

3. The principal chooses to make payments $\tilde{p}_{i,t} \in \mathbb{R}$ to agent i.[15] The winning agent can refuse and keep v. A losing agent can refuse and keep 0.

We first suppose each agent has zero wealth, so that the agents' utilities must be positive in each history of the game. The following result shows that any generalised contract can then be reformulated as a contract where the principal only pays the winner, as in Sections 2 –5.

Proposition 8. *Suppose each agent has zero wealth. Then for any (agent–)self–enforcing contract $\langle \tilde{Q}_{i,t}, \tilde{p}_{i,t} \rangle$, there exists an (agent–)self–enforcing contract $\langle Q_{i,t}, p_t \rangle$ that attains the same time–0 payoffs.*

Proof. See Appendix A.8. □

Intuitively, in order to extract an agent's rents, the principal needs to demand payment before the time of first trade. However, such a payment is impossible if the agent has no wealth. To illustrate, Kaufmann and Lafontaine (1994) show that an average McDonald's franchisee makes ex–post rents of \$400K, but only faces a franchise fee of \$20K. They argue that, since franchisees are expected to actively manage their restaurant, liquidity constraints prevent McDonald's expropriating more of the rent.

A second problem with the full–extraction contracts is that they are not robust to free entry into the market for principals. We say a contract is *cowboy–proof* if a new principal, with costs $c_{i,t} = \infty$, cannot make a strictly positive profit.

Proposition 9. *For any cowboy–proof (agent–)self–enforcing contract $\langle \tilde{Q}_{i,t}, \tilde{p}_{i,t} \rangle$, there exists an (agent–)self–enforcing contract $\langle Q_{i,t}, p_t \rangle$ that attains the same time–0 payoffs.*

Proof. See Appendix A.9. □

Proposition 9 says that if we want the contract to be robust to free entry into the market for principals, then we can limit ourselves to contracts where payments only occur when trade takes place. Thus expanding the space of possible payments does not benefit the principal.[16] To illustrate, McMillan and Woodruff (1999, Table 1) find that only 35% of customer relationships involve any advance payment. One reason is that customers are concerned about cowboys: 6% report that a supplier had failed to deliver goods without returning an advance payment.

6.2 Loyalty without Rents

This paper has argued that loyalty results from the principal trying to minimising agents' rents. However, even if rents can be extracted, the optimal contract may still exhibit loyalty, albeit of

[15] If $\tilde{p}_{i,t} < 0$, then the payment is from agent i to the principal.

[16] This, of course, ignores issues of incentive compatibility. It is worth noting that the employment contract is cowboy–proof and obeys the zero wealth constraint.

a different form. In this case, the principal chooses a generalised relational contract $\langle \tilde{Q}_{i,t}, \tilde{p}_{i,t} \rangle$ to maximise welfare

$$W_0 = E_0 \left[\sum_{t=1}^{\infty} \sum_i \delta^t (v - c_{i,t}) \tilde{Q}_{i,t} \right]$$

subject to her dynamic enforcement constraint,

$$(W_{i,t} - v)\tilde{Q}_{i,t} \geq 0 \qquad (\textbf{??})$$

In addition, suppose costs $c_{i,t}$ are IID.

The principal would thus like to invest in the cheapest agent each period. For fixed N, this satisfies (**??**) if $\delta \geq \hat{\delta}_N$. However, for fixed δ, there is always a large enough N such that (**??**) is not satisfied. That is, the analogue to Proposition 5 is false.

When the welfare–maximising investment rule is not feasible, the optimal contract will induce loyalty. Intuitively, when the principal uses multiple sources, the frequency of trade is reduced and defection becomes more likely. To be more precise, one would expect the principal to reward an agent she uses with cash today and a greater continuation value tomorrow, as in Thomas and Worrall (1988) and Kocherlakota (1996). As a result, the principal would wish to bias trade towards agents with whom she has traded more recently. While a full characterisation of this problem is beyond the scope of the current paper, this form of loyalty is qualitatively different from the loyalty exhibited in Proposition 1, whereby all insiders are treated equally.

7 Conclusion

This paper derives the optimal way for a firm to procure under incomplete contracts. It shows that ongoing relationships can be used to mitigate holdup and analyses the trade–off between maintaining such a long–term relationship and exploiting the gains from trade. We show the optimal contract has a simple characterisation: the principal acts as if there is a fixed cost of initiating a new relationship and biases trade towards 'insiders', with whom she has traded previously, at the expense of 'outsiders'.

The model has many empirical implications. It predicts that a principal will be loyal to her agents, and this loyalty has a specific pattern: she treats all previous trading partners equally; she is biased against new agents; and when adopting new agents, she prefers agents who are large and stable. Looking across economies, loyalty is higher when parties are impatient, when the country has a poorer legal system, and when the principal is purchasing custom made goods. In addition, a principal who exhibits less loyalty receives poorer quality services and is expropriated more frequently. In the optimal contract, suppliers receive rents and prices are sufficiently backloaded to prevent holdup, but not too backloaded to tempt the principal to defect. As a result, prices will tend to increase as the expected frequency of trade decreases. Finally, when a firm has problems committing to a contract the model predicts that, if parties are sufficiently impatient or the value of trade decreases over time, then there may be no trade or radically different trade patterns from that under the optimal contract.

The model can be extended in a number of ways. First, in the current formulation, spot–market contracts yield negative profits and are therefore infeasible. If spot markets were more effective, the principal could foster relationships with some agents, while using others on a temporary basis, as in the Japanese electric machinery industry (Asanuma (1989)). Second, if the principal could choose the level of investment as well as the direction, she would wish to increase her investment over time, as in Thomas and Worrall (1994). As a consequence there would be different levels of insiders, corresponding to the number of times an agent has been used in the past.

The model can also help us think about a number of broader points. First, as highlighted by Kelman (1990), a relational contract requires subjective judgements and therefore cannot be implemented without managerial discretion. This must be traded off against the cost of discretion, such as self–interested managers and possible corruption. Second, the model helps us understand the benefit from writing more complete contracts or integrating production within the firm. For example, it illustrates that formal contracts are most important where there are many potential partners, and where the efficient trading partner changes frequently. Finally, the model raises the issue of how to initiate such a relational contract. To illustrate, GM spent much of the last 20 years trying to implement the Japanese system and build bonds with their suppliers. However, as argued by Liker and Choi (2004), these changes were superficial and did not alter the fundamental nature of their relationship.

A Omitted Proofs

A.1 Proof of Proposition 1

(a) Fix t and suppose that $i \in \mathcal{I}_t$ obeys (3.5). By the principal of optimality, the principal wishes to maximise

$$\Pi_t = E_t\left[\sum_{s=t}^{\infty}\sum_i \delta^{s-t}(v - c_{i,s})Q_{i,s}\right] - E_t\left[\sum_{k \notin \mathcal{I}_t} \delta^{\tau_k(t)-t} v\right] \tag{A.1}$$

Using equation (3.5), the principal prefers trading to both abstaining or trading with another agent, j.

(b)(i) Fix t and suppose $i \notin \mathcal{I}_t$ obeys (3.6a). Fix an allocation $\langle Q_{i,t}\rangle$ such that the principal trades with agent i in period t. Using (A.1), profit is given by

$$\begin{aligned}
\Pi_t(i) &= -c_{i,t} + E_t\left[\sum_{s=t+1}^{\infty}\sum_i \delta^{s-t}(v - c_{i,s})Q_{i,s}\right] - E_t\left[\sum_{k \notin \mathcal{I}_t \setminus \{i\}} \delta^{\tau_k(t+1)-t} v\right] \\
&< -\delta v + E_t\left[\sum_{s=t+1}^{\infty}\sum_i \delta^{s-t}(v - c_{i,s})Q_{i,s}\right] - E_t\left[\sum_{k \notin \mathcal{I}_t \setminus \{i\}} \delta^{\tau_k(t+1)-t} v\right] \\
&\leq 0 + E_t\left[\sum_{s=t+1}^{\infty}\sum_i \delta^{s-t}(v - c_{i,s})Q_{i,s}\right] - E_t\left[\sum_{k \notin \mathcal{I}_t} \delta^{\tau_k(t+1)-t} v\right] \\
&\leq \Pi_t(\emptyset)
\end{aligned}$$

where the first inequality comes from (3.6a) and $\Pi_t(\emptyset)$ is the profit obtained by following the original plan, but abstaining from trade at period t. The proof for (b)(ii) is analogous: In this case, trade with agent i is dominated by trade with agent j.

A.2 Proof of Proposition 2

Assume the distribution of $c_{1,t}$ is less than the distribution of $c_{2,t}$ in terms of first–order stochastic dominance. We will show that whenever $c_{1,t} = c_{2,t}$, then the principal (weakly) prefers agent 1 over agent 2. Since costs are IID, we will henceforth drop time subscripts.

We proceed by induction. First, suppose there are $N - 2$ insiders, which we number $\{3, 4, \ldots, N\}$. The continuation profits when either 1 or 2 are also insiders are given by

$$\Pi_1 = E\big[\max\{v - c_1 + \delta\Pi_1, -c_2 + \delta\Pi_{1,2}, v - c_I + \delta\Pi_1, \delta\Pi_1\}\big] \tag{A.2}$$

$$\Pi_2 = E\big[\max\{-c_1 + \delta\Pi_{1,2}, v - c_2 + \delta\Pi_2, v - c_I + \delta\Pi_2, \delta\Pi_2\}\big] \tag{A.3}$$

where c_I is the lowest cost of $\{3,4,\ldots,N\}$, and $\Pi_{1,2}$ are the profits when both 1 and 2 are insiders.

By contradiction, assume $\Pi_2 > \Pi_1$. Then (A.2) implies

$$\Pi_1 \geq E\big[\max\{v - c_1 + \delta\Pi_2, -c_2 + \delta\Pi_{1,2}, v - c_I + \delta\Pi_2, \delta\Pi_2\}\big] - \delta(\Pi_2 - \Pi_1) \tag{A.4}$$

Lemma 3 (Shaked and Shanthikumar (2007, Thm 1.A.10))**.** *Suppose (c_1, c_2) are independent variables and c_2 first–order stochastically dominates c_1. Suppose $\phi_1(c_1,c_2)$ and $\phi_2(c_1,c_2)$ satisfy the following conditions:*
(a) $\phi_1(c_1,c_2) - \phi_2(c_1,c_2)$ is decreasing in c_1.
(b) $\phi_1(c_1,c_2) - \phi_2(c_1,c_2)$ is increasing in c_2.
(c) $\phi_1(c_1,c_2) = \phi_2(c_2,c_1)$.
Then $E[\phi_1(c_1,c_2)] - E[\phi_2(c_1,c_2)] \geq 0$.

Fixing c_I, let $\phi_1(c_1,c_2)$ be the term in the brackets on right hand side of (A.4) and $\phi_2(c_1,c_2)$ be the term in the brackets on right hand side of (A.3). We can now apply Lemma 3, where conditions (a) and (b) follow from the fact than agent $i \in \{1,2\}$ is more likely to be used when he is an insider than when he is an outsider. Taking expectations over c_I,

$$\Pi_2 - \Pi_1 \leq E[\phi_2] - E[\phi_1] + \delta(\Pi_2 - \Pi_1) \leq \delta(\Pi_2 - \Pi_1) \tag{A.5}$$

Hence $\Pi_1 \geq \Pi_2$, yielding a contradiction.

Proceeding by induction, suppose there is a group of insiders I, excluding agents 1 and 2. The continuation profits when either 1 or 2 are also insiders are given by

$$\Pi_1 = E\big[\max\{v - c_1 + \delta\Pi_1, -c_2 + \delta\Pi_{1,2}, v - c_I + \delta\Pi_1, v - c_O + \delta\Pi_{1,O}, \delta\Pi_1\}\big] \tag{A.6}$$

$$\Pi_2 = E\big[\max\{-c_1 + \delta\Pi_{1,2}, v - c_2 + \delta\Pi_2, v - c_I + \delta\Pi_2, v - c_O + \delta\Pi_{2,O}, \delta\Pi_2\}\big] \tag{A.7}$$

where c_I is the lowest cost of the insiders (excluding 1 and 2), c_O is the lowest cost of the outsiders (excluding 1 and 2), $\Pi_{1,2}$ are the profits when both 1 and 2 are insiders, and $\Pi_{1,O}$ are the profits when both 1 and one outsider are insiders.

By induction, we know that $\Pi_{1,O} \geq \Pi_{2,O}$. By contradiction, assume $\Pi_2 > \Pi_1$. Then (A.6) implies

$$\Pi_1 \geq E\big[\max\{v - c_1 + \delta\Pi_2, -c_2 + \delta\Pi_{1,2}, v - c_I + \delta\Pi_2, v - c_O + \delta\Pi_{2,O}, \delta\Pi_2\}\big] - \delta(\Pi_2 - \Pi_1) \tag{A.8}$$

Fixing (c_I, c_O), let $\phi_1(c_1,c_2)$ be the term in the brackets on the right hand side of (A.8) and $\phi_2(c_1,c_2)$ be the term in the brackets on the right hand side of (A.7). Taking expectations over

(c_I, c_O), Lemma 3 yields equation (A.5). Hence $\Pi_1 \geq \Pi_2$, providing a contradiction.

Finally, suppose there is some set of insiders $\mathcal{I}$ that does not include either 1 or 2. The principal's profits are thus

$$\Pi = E\big[\max\{-c_1 + \delta\Pi_1, -c_2 + \delta\Pi_2, v - c_I + \delta\Pi, v - c_O + \delta\Pi_O, \delta\Pi\}\big]$$

where Π_1 are the profits when 1 becomes an insider, Π_2 the profits when 2 becomes an insider, and Π_O are the profits when the lowest cost outsider (excluding 1 and 2) becomes an insider. Since $\Pi_1 \geq \Pi_2$, the principal prefers agent 1 over agent 2, as required.

A.3 Proof of Proposition 4

The optimal cutoff is given by equation (3.13). We thus seek to show that $\Delta\Pi_{n+1} := \Pi_{n+1} - \Pi_n$ decreases in n. First, let us write

$$\Pi_{n+2} = E\big[\max\{v - c_1 + \delta\Pi_{n+2}, v - c_2 + \delta\Pi_{n+2}, v - c_I + \delta\Pi_{n+2}, -c_O + \delta\Pi_{n+3}, \delta\Pi_{n+2}\}\big] \quad \text{(A.9)}$$
$$\Pi_{n+1} = E\big[\max\{-c_1 + \delta\Pi_{n+2}, v - c_2 + \delta\Pi_{n+1}, v - c_I + \delta\Pi_{n+1}, -c_O + \delta\Pi_{n+2}, \delta\Pi_{n+1}\}\big] \quad \text{(A.10)}$$

where c_1 and c_2 are the costs of agents 1 and 2, c_I is the lowest insider cost (excluding agents 1 and 2), and c_O is the lowest outsider cost (excluding agents 1 and 2). Similarly, it is useful to write

$$\Pi_{n+1} = E\big[\max\{v - c_1 + \delta\Pi_{n+1}, -c_2 + \delta\Pi_{n+2}, v - c_I + \delta\Pi_{n+1}, -c_O + \delta\Pi_{n+2}, \delta\Pi_{n+1}\}\big] \quad \text{(A.11)}$$
$$\Pi_n = E\big[\max\{-c_1 + \delta\Pi_{n+1}, -c_2 + \delta\Pi_{n+1}, v - c_I + \delta\Pi_n, -c_O + \delta\Pi_{n+1}, \delta\Pi_n\}\big] \quad \text{(A.12)}$$

Subtracting $\delta\Pi_{n+2}$ from (A.9) and (A.10),

$$\begin{aligned}\Delta\Pi_{n+2} = {} & E\big[\max\{v - c_1, v - c_2, v - c_I, -c_O + \delta\Delta\Pi_{n+3}, 0\}\big] \\ & - E\big[\max\{-c_1, v - c_2 - \delta\Delta\Pi_{n+2}, v - c_I - \delta\Delta\Pi_{n+2}, -c_O, -\delta\Delta\Pi_{n+2}\}\big]\end{aligned} \quad \text{(A.13)}$$

Similarly, subtracting $\delta\Pi_{n+1}$ from (A.11) and (A.10),

$$\begin{aligned}\Delta\Pi_{n+1} = {} & E\big[\max\{v - c_1, -c_2 + \delta\Delta\Pi_{n+2}, v - c_I, -c_O + \delta\Delta\Pi_{n+2}, 0\}\big] \\ & - E\big[\max\{-c_1, -c_2, v - c_I - \delta\Delta\Pi_{n+1}, -c_O, -\delta\Delta\Pi_{n+1}\}\big]\end{aligned} \quad \text{(A.14)}$$

We now prove the result by induction. Since $\underline{c} > 0$ or $v > \overline{c}$, equation (3.10) implies $n^* < \infty$. Let $n = \min\{N - 2, n^*\}$, allowing us to drop the c_O term in equations (A.13) and (A.14). By contradiction, assume that $\Delta\Pi_{n+2} > \Delta\Pi_{n+1}$.

First, suppose $c_1 \leq c_2$. Then (A.13) becomes

$$\Delta\Pi_{n+2} \leq E\big[\max\{v - c_1, v - c_I, 0\}\big] - E\big[\max\{-c_1, v - c_I - \delta\Delta\Pi_{n+2}, -\delta\Delta\Pi_{n+2}\}\big]$$

where we use the "max" operator. Equation (A.14) becomes

$$\begin{aligned}\Delta\Pi_{n+1} &= E\big[\max\{v - c_1, v - c_I, 0\}\big] - E\big[\max\{-c_1, v - c_I - \delta\Delta\Pi_{n+1}, -\delta\Delta\Pi_{n+1}\}\big]\\ &\geq E\big[\max\{v - c_1, v - c_I, 0\}\big] - E\big[\max\{-c_1, v - c_I - \delta\Delta\Pi_{n+2}, -\delta\Delta\Pi_{n+2}\}\big] - \delta(\Delta\Pi_{n+2} - \Delta\Pi_{n+1})\end{aligned}$$

where the first line uses $\delta\Delta\Pi_{n+2} \leq v$ and the second uses $\Delta\Pi_{n+2} > \Delta\Pi_{n+1}$. Comparing these equations, we have $\Delta\Pi_{n+2} - \Delta\Pi_{n+1} \leq \delta(\Delta\Pi_{n+2} - \Delta\Pi_{n+1})$, implying that $\Delta\Pi_{n+1} \geq \Delta\Pi_{n+2}$ and yielding a contradiction.

Second, assume $c_1 \geq c_2$. Equation (A.13) becomes

$$\begin{aligned}\Delta\Pi_{n+2} &= E\big[\max\{v - c_2, v - c_I, 0\}\big] - E\big[\max\{v - c_2 - \delta\Delta\Pi_{n+2}, v - c_I - \delta\Delta\Pi_{n+2}, -\delta\Delta\Pi_{n+2}\}\big]\\ &= \delta\Delta\Pi_{n+2}\end{aligned}$$

Similarly, equation (A.14) becomes

$$\begin{aligned}\Delta\Pi_{n+1} &\geq E\big[\max\{-c_2 + \delta\Delta\Pi_{n+2}, v - c_I, 0\}\big] - E\big[\max\{-c_2, v - c_I - \delta\Delta\Pi_{n+1}, -\delta\Delta\Pi_{n+1}\}\big]\\ &\geq E\big[\max\{-c_2 + \delta\Delta\Pi_{n+2}, v - c_I, 0\}\big] - E\big[\max\{-c_2, v - c_I - \delta\Delta\Pi_{n+2}, -\delta\Delta\Pi_{n+2}\}\big] - \delta(\Delta\Pi_{n+2} - \Delta\Pi_{n+1})\\ &= \delta\Delta\Pi_{n+1}\end{aligned}$$

where the first line uses the "max" operator, and the second uses $\Delta\Pi_{n+2} > \Delta\Pi_{n+1}$. Hence we have $\Delta\Pi_{n+1} \geq \Delta\Pi_{n+2}$, yielding a contradiction.

Continuing by induction, consider $n < \min\{n^*, N-2\}$. By contradiction, assume $\Delta\Pi_{n+2} > \Delta\Pi_{n+1}$. First, assume $c_1 \leq c_2$. Then (A.13) becomes

$$\Delta\Pi_{n+2} \leq E\big[\max\{v - c_1, v - c_I, -c_O + \delta\Delta\Pi_{n+3}, 0\}\big] - E\big[\max\{-c_1, v - c_I - \delta\Delta\Pi_{n+2}, -c_O, -\delta\Delta\Pi_{n+2}\}\big]$$

where we use the max operator. Equation (A.14) becomes

$$\begin{aligned}\Delta\Pi_{n+1} &= E\big[\max\{v - c_1, v - c_I, -c_O + \delta\Delta\Pi_{n+2}, 0\}\big] - E\big[\max\{-c_1, v - c_I - \delta\Delta\Pi_{n+1}, -c_O, -\delta\Delta\Pi_{n+1}\}\big]\\ &\geq E\big[\max\{v - c_1, v - c_I, -c_O + \delta\Delta\Pi_{n+3}, 0\}\big] - E\big[\max\{-c_1, v - c_I - \delta\Delta\Pi_{n+2}, -c_O, -\delta\Delta\Pi_{n+2}\}\big]\\ &\quad - \delta(\Delta\Pi_{n+2} - \Delta\Pi_{n+1})\end{aligned}$$

where the first line uses $\delta\Delta\Pi_{n+2} \leq v$, and the second uses $\Delta\Pi_{n+2} \geq \Delta\Pi_{n+3}$ from the induction hypothesis and the assumption that $\Delta\Pi_{n+2} > \Delta\Pi_{n+1}$. Hence we have $\Delta\Pi_{n+1} \geq \Delta\Pi_{n+2}$, yielding a contradiction.

Second, assume $c_1 \geq c_2$. Equation (A.13) becomes

$$\begin{aligned}\Delta\Pi_{n+2} &= E\big[\max\{v - c_2, v - c_I, -c_O + \delta\Delta\Pi_{n+3}, 0\}\big] - E\big[\max\{v - c_2 - \delta\Delta\Pi_{n+2}, v - c_I - \delta\Delta\Pi_{n+2}, -c_O, -\delta\Delta\Pi_{n+2}\}\big] \\ &\leq \delta\Delta\Pi_{n+2}\end{aligned}$$

where the first line uses $c_1 \geq c_2$, and the second line uses the induction hypothesis. Similarly, equation (A.14) becomes

$$\begin{aligned}\Delta\Pi_{n+1} &\geq E\big[\max\{-c_2 + \delta\Delta\Pi_{n+2}, v - c_I, -c_O + \delta\Delta\Pi_{n+2}, 0\}\big] - E\big[\max\{-c_2, v - c_I - \delta\Delta\Pi_{n+1}, -c_O, -\delta\Delta\Pi_{n+1}\}\big] \\ &\geq E\big[\max\{-c_2 + \delta\Delta\Pi_{n+2}, v - c_I, -c_O + \delta\Delta\Pi_{n+2}, 0\}\big] - E\big[\max\{-c_2, v - c_I - \delta\Delta\Pi_{n+2}, -c_O, -\delta\Delta\Pi_{n+2}\}\big] \\ &\qquad - \delta(\Delta\Pi_{n+2} - \Delta\Pi_{n+1}) \\ &= \delta\Delta\Pi_{n+1}\end{aligned}$$

where the first line uses the "max" operator and the second line uses the assumption that $\Delta\Pi_{n+2} > \Delta\Pi_{n+1}$. Hence we have $\Delta\Pi_{n+1} \geq \Delta\Pi_{n+2}$, yielding a contradiction.

A.4 Proof of Proposition 5

We seek to show that, for δ sufficiently high, (DEP$_{\mathrm{B}}$) is slack under the optimal contract. This implies that the (DEP$_{\mathrm{M}}$) constraint is also slack. If $Q_{i,t} = 1$, the profit from relationship i is

$$\begin{aligned}\Pi_{i,t} &= -c_{i,t} + \delta E_t[W_{i,t+1}] \\ &\geq -v + \frac{1}{n^*}\frac{\delta}{1-\delta}\omega(n^*)\end{aligned} \tag{A.15}$$

where n^* is the maximum number of agents, and $\omega(n) := E[\max\{v - c_{1:n}, 0\}]$ is welfare when using n agents. Let us define $\omega(0) := 0$. Two facts are worth noting. First, $\omega(n)$ is increasing in n and converges to $E[v - \underline{c}]$. Second, the marginal welfare of an extra agent, $\Delta\omega(n) := \omega(n) - \omega(n-1)$ decreases in n and converges to zero.

The proof rests on two lemmas.

Lemma 4. *For any increasing sequence of integers $\{n_i\}$,*

$$\sum_{i\geq 1} \frac{n_i - n_{i-1}}{n_i} = \infty$$

Proof. Pick an infinite subsequence of the integers $\{m_j\}$ as follows. Let $m_1 := n_1$ and $m_j =$

$\min\{n_i : n_i \geq 2m_{j-1}\}$. Then

$$\begin{aligned}\sum_{i\geq 1}\frac{n_i - n_{i-1}}{n_i} &= \sum_j \sum_{\{i:\, m_j \geq n_i > m_{j-1}\}} \frac{n_i - n_{i-1}}{n_i}\\ &\geq \sum_j \sum_{\{i:\, m_j \geq n_i > m_{j-1}\}} \frac{n_i - n_{i-1}}{m_j}\\ &= \sum_j \frac{m_j - m_{j-1}}{m_j}\end{aligned}$$

By construction, $\frac{m_j - m_{j-1}}{m_j} \geq \frac{1}{2}$. Hence the sum is infinite. □

Lemma 5. *$n\Delta\omega(n) \to 0$ as $n \to \infty$*

Proof. Since $\Delta\omega(n) > 0$ $(\forall n)$, $\liminf n\Delta\omega(n) \geq 0$. By contradiction, suppose that $\limsup n\Delta\omega(n) = k > 0$. Then there exists a subsequence of integers $\{n_i\}$ such that

$$n_i\Delta\omega(n_i) \geq k - \epsilon > 0 \qquad (\forall i) \tag{A.16}$$

Abusing notation, let $n_i(n) = \min\{n_i : n_i \geq n\}$ be the next integer in the subsequence after an arbitrary integer n. We now obtain the following contradiction,

$$\begin{aligned} v - \underline{c} &= \sum_{n\geq 1}\Delta\omega(n)\\ &\geq \sum_{n\geq 1}\Delta\omega(n_i(n))\\ &= \sum_{i\geq 1}(n_i - n_{i-1})\Delta\omega(n_i)\\ &\geq (k-\epsilon)\sum_{i\geq 1}\frac{n_i - n_{i-1}}{n_i}\\ &= \infty\end{aligned}$$

The first line follows from $\omega(1) = E[v - c]$ and $\lim\omega(n) = v - \underline{c}$. The second uses the fact that $\Delta\omega(n)$ is decreasing. The fourth line uses (A.16), while the fifth line follows from Lemma 4. □

Using equation (3.9), we can place an upper bound on the number of insiders, n^*.[17] Substituting (3.9) into (A.15),

$$\Pi_{i,t} \geq -v + \frac{\omega(n^*)}{n^*\Delta\omega(n^*)}\left[\max\{v - \bar{c}, 0\} + \underline{c}\right] \tag{A.17}$$

[17]Of course, this bound will not be reached if $n^* > N$.

Since either $\underline{c} > 0$ or $v > \overline{c}$, then $\max\{v - \overline{c}, 0\} + \underline{c} > 0$. As $\delta \to 1$, so $n^*(\delta)$ increases monotonically without bound. Consequently, $\omega(n^*) \to v - \underline{c}$ and, by Lemma 5, $n^*[\Delta\omega(n^*)] \to 0$. Hence $\Pi_{i,t} \to \infty$, as required. Observe that this argument applies for any value of N, including $N = \infty$.

A.5 Derivation of Example 2

Suppose $c_{i,t} \sim U[0,1]$ and $v > 1$. As in equation (A.15), profits at time t are

$$\begin{aligned} \Pi_{i,t} &= -c_{i,t} + \delta E_t[W_{i,t+1}] \\ &\geq -1 + \frac{1}{n^*}\frac{\delta}{1-\delta}\omega(n^*) \\ &\geq -1 + \left[\frac{1}{n^*(n^*+1)}\right]^{1/2} \frac{\delta}{1-\delta}(v-1) \end{aligned} \tag{A.18}$$

since $\omega(n) = v - \frac{1}{1+n}$. Using equation (3.9),

$$\frac{\delta}{1-\delta}\frac{1}{n^*(n^*+1)} \geq (v-1) \tag{A.19}$$

Substituting (A.19) into (A.18) yields,

$$\Pi_{i,t} \geq -1 + \left[\frac{\delta}{1-\delta}\right]^{1/2} (v-1)^{3/2}$$

Hence $\Pi_{i,t} \geq 0$ if $(v-1)^3 \geq (1-\delta)/\delta$. That is,

$$\delta \geq \frac{1}{1+(v-1)^3}$$

A.6 Proof of Proposition ??

The problem is to maximise (**??**) subject to (**??**).[18] First, suppose that the agent is an insider (i.e. that they have bought previously). By the principle of optimality, the principal wishes to maximise

$$W_t = (v - c_t)Q_t + \delta E_t[W_{t+1}]$$

subject to

$$\big(\delta E_t[W_{t+1}] - c_t\big)Q_t \geq 0 \tag{??}$$

[18] Since there is only one agent, we drop i subscripts.

By pointwise optimisation, the solution is monotone: that is, trade occurs if c_t is below some threshold. Since the problem is stationary, the value function, $W(\kappa)$, can thus be written as

$$E[W(\kappa)] := \frac{1}{1-\delta} E[(v-c)\mathbf{1}_{c\leq\kappa}]$$

Then κ^{**} is characterised by the largest cost that satisfies (**??**),

$$\kappa^{**} = \max\{\kappa \leq v \,:\, \delta E[W(\kappa)] - \kappa \geq 0\} \tag{A.20}$$

Initial profit $\Pi(\kappa^*)$ is defined by

$$\begin{aligned}\Pi(\kappa^*) &= \int_{\underline{c}}^{\kappa^*} (\delta E[W(\kappa^{**})] - c)\, dF(c) + \int_{\kappa^*}^{\overline{c}} \delta\Pi(\kappa^*) dF(c) \\ &= \frac{\int_{\underline{c}}^{\kappa^*} (\delta E[W(\kappa^{**})] - c)\, dF(c)}{1 - \delta(1 - F(\kappa^*))}\end{aligned} \tag{A.21}$$

The initial cutoff κ^* is chosen to maximise $\Pi(\kappa^*)$ subject to (**??**). Since (**??**) must hold, $\kappa^* \leq \kappa^{**}$. Finally, we wish to show that $\kappa^* \leq c^*$ and $\kappa^{**} \leq v$. The latter follows from (A.20). To show the former note that $\Pi(\kappa^*)$ is log–supermodular in $(\kappa^*, E[W])$, so that κ^* is increasing in $E[W]$. Since expected welfare $E[W]$ is higher under the optimal ASE contract, the first period cutoff is higher, as required.

A.7 Proof of Proposition 6

First, we show the contract is agent–self–enforcing. Under the employment contract, if $Q_{i,t} = 1$, agent i anticipates future rents v. Hence (DEA) holds, and the agent has no incentive to defect.

Second, we show incentive compatibility. At time t, the principal's profit is

$$\Pi_t = E_t\left[\sum_{s=t}^{\infty} \delta^{s-t}(v - c_{I_t,t})\right] - E_t\left[\sum_{i\notin\mathcal{I}_t} \delta^{\tau_i(t)-t} v\right] - \sum_{i\in\mathcal{I}_t} v \tag{A.22}$$

Observe that the last term in (A.22) is sunk and can be ignored. The investment plan, $Q_{i,t}$, chosen to maximise time–0 profits (3.4) thus maximises time–t profits (A.22) by the principle of optimality. The principal therefore cannot gain by lying about her costs, altering the investment plan.

Third, the contract is self–enforcing if the principal has no incentive to deviate. When $Q_{i,t} = 1$ the principal must be willing to invest in i. That is,

$$\Pi_{i,t} = W_{i,t} - v \geq 0 \tag{A.23}$$

When $Q_{i,t} = 0$ and $i \in \mathcal{I}_t$, the principal must be willing to pay the agent $p_{i,t} = (1-\delta)v$. That is,

$$\Pi_{i,t} = W_{i,t} - v \geq 0 \tag{A.24}$$

When $Q_{i,t} = 0$ and $i \notin \mathcal{I}_t$, the principal cannot defect. Putting (A.23) and (A.24) together, yields (DEP$^{\text{EC}}$).

A.8 Proof of Proposition 8

Pick a contract $\langle \tilde{Q}_{i,t}, \tilde{p}_{i,t} \rangle$. Define utility of agent i at time t by

$$\tilde{U}_{i,t} = E_t \left[\sum_{s=t}^{\infty} \delta^{s-t} \tilde{p}_{i,s} \right]$$

Similarly, define the profit from relationship i at time t by,

$$\tilde{\Pi}_{i,t} = E_t \left[\sum_{s=t}^{\infty} \delta^{s-t} \left((v - c_{i,s})\tilde{Q}_{i,s} - \tilde{p}_{i,s} \right) \right]$$

Denote the times the principal invests in i by $\{t_1^i, t_2^i, \ldots\}$. In terms of our previous notation, $t_1^i = \tau_i(0)$ and $t_{n+1}^i = \tau_i(t_n^i + 1)$ for $n \geq 1$.

We now establish three facts. First, observe that the principal can always defect on i at time t_n^i by refusing to invest. This would yield the agent utility

$$\hat{U}_{i,n} := \sum_{s=1}^{t_n^i - 1} \delta^s \tilde{p}_{i,s} \tag{A.25}$$

The zero wealth constraint thus says that (A.25) is positive for each t_n^i.

Second, when $Q_{i,t} = 1$, the agent can defect by refusing to pay the principal, keeping v. Hence the relational contract must satisfy,

$$\tilde{U}_{i,t} \geq v \tag{A.26}$$

Third, when $Q_{i,t} = 1$, the principal can defect by refusing to invest. Hence the relational contract must satisfy,

$$\tilde{\Pi}_{i,t} \geq 0 \tag{A.27}$$

We now construct a new contract, $\langle Q_{i,t}, p_t \rangle$, where payments are voluntary and only occur when trade takes place. Let $Q_{i,t} = \tilde{Q}_{i,t}$ and

$$
\begin{aligned}
p_t &= E_t\left[\sum_{s=t}^{t_2^i-1} \delta^{s-t}\tilde{p}_{i,s}\right] + \sum_{s=1}^{t_1^i-1} \delta^{s-t}\tilde{p}_{i,s} && \text{for } t = t_1^i && \text{(A.28)}\\
&= E_t\left[\sum_{s=t}^{t_{n+1}^i-1} \delta^{s-t}\tilde{p}_{i,s}\right] && \text{for } t = t_n^i \text{ and } n > 1
\end{aligned}
$$

Let $U_{i,t}$ and $\Pi_{i,t}$ be the utilities and profits from this new contract.[19] The payoffs from the two contracts are related as follows. At time 0, the contracts yield the same payoffs. That is, $U_{i,0} = \tilde{U}_{i,0}$ and $\Pi_{i,0} = \tilde{\Pi}_{i,0}$. At time $t = t_1^i$, we have

$$U_{i,t} = \delta^{-t}\tilde{U}_{i,0} = \tilde{U}_{i,t} + \delta^{-t}\hat{U}_{i,1} \tag{A.29}$$

$$\Pi_{i,t} = \delta^{-t}\tilde{\Pi}_{i,0} = \tilde{\Pi}_{i,t} - \delta^{-t}\hat{U}_{i,1} \tag{A.30}$$

At time $t = t_n^i$ and $n > 1$, we have

$$U_{i,t} = \tilde{U}_{i,t} \tag{A.31}$$

$$\Pi_{i,t} = \tilde{\Pi}_{i,t} \tag{A.32}$$

We now make three claims. First, the new contract satisfies the agent's wealth constraint. Equation (A.25) implies that for each n,

$$\sum_{m=1}^{n-1} \delta^{t_m^i} p_m = \sum_{s=1}^{t_1^i-1} \delta^s \tilde{p}_{i,s} + \sum_{m=1}^{n-1} E_{t_m^i}\left[\sum_{s=t_m^i}^{t_{m+1}^i-1} \delta^s \tilde{p}_{i,s}\right]$$

is positive. Hence the agent's wealth is positive in every history.

Second, the agent's dynamic enforcement constraint is satisfied. In period $t = t_1^i$, equations (A.29) and (A.26) and the fact that $\hat{U}_{i,1} \geq 0$, implies $U_{i,t} \geq v$. In subsequent periods of trade, equations (A.26) and (A.31) imply that $U_{i,t} \geq v$. Hence the new contract is ASE when the original contract is ASE.

Third, the principal's dynamic enforcement constraint is satisfied. In period $t = t_1^i$, equation (A.30) and the fact that $\tilde{\Pi}_{i,0} \geq 0$ implies that $\Pi_{i,t} \geq 0$. In subsequent periods of trade, equations (A.27) and (A.32) imply that $\Pi_{i,t} \geq 0$. Hence the new contract is SE when the original contract is SE.

[19] Under this construction, we have $p_t \in \mathbb{R}$. Under the principal's optimal contract, we can restrict prices to $p_t \in [0, v]$, as shown in Section 3.2.

A.9 Proof of Proposition 9

Pick a contract $\langle \tilde{Q}_{i,t}, \tilde{p}_{i,t} \rangle$. Cowboy–proofness says that a cowboy's profits from agent i,

$$\hat{\Pi}_i = - \sum_{t=1}^{\tau_i(0)-1} \delta^t \tilde{p}_{i,t}$$

must be negative. This says that $\hat{U}_{i,1} \geq 0$, as defined by (A.25). The new contract $\langle Q_{i,t}, p_t \rangle$ is constructed by setting $Q_{i,t} = \tilde{Q}_{i,t}$ and defining prices by (A.28). Verifying this new contract satisfies the dynamic enforcement constraints is then the same as Proposition 8.

References

Abreu, D. (1988): "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica*, 56(2), 383–396.

Albuquerque, R., and H. Hopenhayn (2004): "Optimal Lending Contracts and Firm Dynamics," *Review of Economic Studies*, 71(2), 285–315.

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

Bewley, T. F. (1999): *Why Wages Don't Fall During a Recession*. Harvard University Press.

Calzolari, G., and G. Spagnolo (2006): "Reputational Commitment and Collusion in Procurement," Working paper, University of Bologna.

DeMarzo, P., and M. Fishman (2007): "Agency and Optimal Investment Dynamics," *Review of Financial Studies*, 20(1), 151–188.

Greif, A. (1993): "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition," *American Economic Review*, 83(3), 525–548.

Hart, O. (1995): *Firms, Contracts and Financial Structure*. Oxford University Press, Oxford, UK.

Hart, O., and J. Moore (1994): "A Theory of Debt Based on the Inalienability of Human Capital," *Quarterly Journal of Economics*, 109(4), 841–879.

——— (1999): "Foundations of Incomplete Contracts," *Review of Economic Studies*, 66(1), 115–138.

Johnson, S., J. McMillan, and C. Woodruff (2002): "Courts and Relational Contracts," *Journal of Law, Economics and Organization*, 18, 221–277.

Kandori, M. (1992): "Social Norms and Community Enforcement," *Review of Economic Studies*, 59(1), 63–80.

Kaufmann, P. J., and F. Lafontaine (1994): "Costs of Control: The Source of Economic Rents for McDonald's Franchisees," *Journal of Law and Economics*, 37(2), 417–453.

Kelman, S. (1990): *Procurement and Public Management*. AEI Press, Washington, D.C.

Klemperer, P. (1995): "Competition when Consumers have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics, and International Trade," *Review of Economic Studies*, 62(4), 515–539.

Kocherlakota, N. R. (1996): "Implications of Efficient Risk Sharing without Commitment," *Review of Economic Studies*, 63(4), 595–609.

Levin, J. (2002): "Multilateral Contracting and the Employment Relationship," *Quarterly Journal of Economics*, 117(3), 1075–1103.

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

Liker, J. K., and T. Y. Choi (2004): "Building Deep Supplier Relationships," *Harvard Business Review*, 82(12), 104–113.

Macaulay, S. (1963): "Non-Contractual Relations in Business: A Preliminary Study," *American Sociological Review*, 28(1), 55–67.

MacLeod, W. B., and J. M. Malcomson (1989): "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment," *Econometrica*, 57(2), 447–480.

Mathewson, G. F., and R. A. Winter (1985): "The Economics of Franchise Contracts," *Journal of Law and Economics*, 28(3), 503–526.

McMillan, J., and C. Woodruff (1999): "Dispute Prevention without Courts in Vietnam," *Journal of Law, Economics and Organization*, 15(3), 637–658.

Milgrom, P., and J. Roberts (1997): "Johnson Controls, Inc.," *HBS Case BE9*.

Ramey, G., and J. Watson (1997): "Contrsctual Fragility, Job Destruction and Business Cycles," *Quarterly Journal of Economics*, 112(3), 873–911.

Sako, M. (1996): "Suppliers Associations in the Japanese Auto Industry: Collective Action for Technology Diffusion?," *Cambridge Journal of Economics*, 20(6), 651–671.

Segal, I. (1999): "Complexity and Renegotiation: A Foundation for Incomplete Contracts," *Review of Economic Studies*, 66(1), 57–82.

Shaked, M., and J. G. Shanthikumar (2007): *Stochastic Orders*. Springer.

Shapiro, C. (1983): "Premiums for High Quality Products as Returns to Reputations," *Quarterly Journal of Economics*, 98(4), 659–680.

Shapiro, C., and J. E. Stiglitz (1984): "Equilibrium Unemployment as a Worker Discipline Device," *American Economic Review*, 74(3), 433–444.

Sigouin, C. (2003): "Investment Decisions, Financial Flows, and Self–Enforcing Contracts," *International Economic Review*, 44(4), 1359–1382.

Sobel, J. (2006): "For Better or Forever: Formal versus Informal Enforcement," *Journal of Labor Economics*, 24(2), 271–297.

Thomas, J., and T. Worrall (1988): "Self-Enforcing Wage Contracts," *Review of Economic Studies*, 55(4), 541–553.

——— (1994): "Foreign Direct Investment and the Risk of Expropriation," *Review of Economic Studies*, 61(1), 81–108.