

Relational Contracts and the Value of Loyalty

Simon Board*

This Version: July 7, 2006

First Version: July 2006.

Abstract

This paper examines how relational contracting affects the pattern of trade across the economy. We suppose a firm (principal) repeatedly chooses among a set of potential trading partners (agents) under the threat of holdup. The possibility of ex-post opportunism allows agents to collect rents, which act like a fixed cost that the principal must pay when initiating a new relationship. The principal responds by dividing agents into “insiders”, with whom she has previously traded, and “outsiders”, with whom she has never traded. If the principal is sufficiently patient, the profit-maximising contract then uses insiders efficiently, while being biased against outsiders. This optimal strategy can be implemented by a “maintenance contract” that is robust to asymmetric information.

1 Introduction

Holdup is one of the most pervasive forces in economics, shaping trading relationships, organisations and even entire societies. In early trading communities, holdup plagued the relationships between merchants and their agents. The mechanisms used to overcome the holdup problem then had enormous effects on these communities’ subsequent development (Greif (1994)). In developing countries, the problem is just as severe since contracts enjoy little legal protection. The threat of holdup then restricts the level of investment, the scope of trade, and ultimately the level of growth (McMillan and Woodruff (1999a, 1999b)). In developed countries, despite the presence of a functioning legal system, holdup is still a problem: 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 circumstances, ongoing relationships are used to mitigate the holdup problem. However, maintaining relationships can reduce of scope of trade. In the words of McMillan and

*Department of Economics, University of Toronto. <http://www.economics.utoronto.ca/board>. I thank the seminar audience at CEMFI’s MadMac for comments.

Woodruff (1999b, p. 1315), “exclusion is the corollary of ongoing relationships.” In a survey of transitional Eastern European countries, Johnson, McMillan, and Woodruff (2002) illustrate the economic significance of loyalty. 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. These switching costs are significantly larger in countries with less effective court systems and for goods that are custom-built.

The purpose of this paper is to provide a model which explains how relational contracts lead to loyalty. This model allows us to derive the exact pattern of trade in the economy and characterise the distortions induced by relational contracts. Our theory is based on a simple observation: agents’ ability to hold up the principal allows them to garner rents. Agents’ care about their lifetime utilities, so these rents are the same whether an agent trades once or trades one hundred times, acting like a fixed cost of initiating a new relationship. In order to minimise these rents, the principal would like to direct trade towards agents she has used in the past. We show that such a bias is time inconsistent but can be implemented if the principal is sufficiently patient. The profit-maximising trade pattern can also be implemented by a “maintenance contract” that is robust to adverse selection.

1.1 Outline of the Paper

We start with a basic holdup model. Each period a principal contracts with one of N agents. In order to make a product the principal must first make a specific investment in the chosen agent. If the principal invests in agent i , then i produces and sells a product worth v . After the sale, the agent chooses how much money to give back to the principal.

We suppose that the cost of the specific investment differs over time and over agents. This cost variation may result from agent i having a backlog of orders at time t , or having relatively little expertise in the product required that period. As a result, the principal’s efficient trading partner will tend to vary over time. We then examine the profit-maximising contract for the principal among the class of contracts that are self-enforcing.

In Section 3, we first suppose that the principal commits to a specified investment strategy. This acts as an initial benchmark and allows 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 large enough rent so that they resist 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 gains from trade against the cost

of starting new relationships. The principal’s profit–maximising contract thus divides agents into two groups: “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, 1994) subsequently observed that an agent who trades more frequently demands lower per–period payments. Grief’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.¹ The current paper uses this observation to explicitly identify the cost associated with a new relationship. We then allow the principal to optimally choose their partner each period, rather than forcing them to pick randomly from a pool.

The profit–maximising 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 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 hold–up 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 profit–maximising 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 a maintenance contract, where insiders are paid a fixed amount per period, independent of whether trade occurs or not. The advantage of the maintenance contract is that, since the payments to insiders are sunk, the principal has no

¹The principal is female, while the agents are male.

incentive to misstate her true costs. The disadvantage is that the principal has more incentive to renege on this contract than the on profit-maximising contract from Section 3. Nevertheless, under IID costs, the maintenance contract is self-enforcing when the principal is sufficiently patient.

Our results rely heavily on the fact that agents' can guarantee themselves rents by threatening to hold up the principal. However, if the principle can use contractible transfers then she can extract all rents from the agents. In Section 6 we discuss the feasibility of such contractible transfers. In particular, we argue that these transfers are not robust to free entry into the market for principals, providing a foundation for our theory.

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 maximises joint surplus and is stationary. Reputation is used to prevent holdup in a number of related areas. For example, Klein and Leffler (1981) and Shapiro (1983) examine a firm's choice of product quality when selling an experience good.

The current paper argues that loyalty results from minimising agents' rents. A companion paper, Board (2006), examines a complimentary explanation for loyalty in relational contracts. In this paper I show that, if agents are relatively impatient, then loyalty results from the inability of the principal to commit to the relational contract. Under the profit-maximising contract, the principal is then more loyal to agents with whom she has traded more recently. This property is qualitatively different from the loyalty exhibited in the current paper, whereby all insiders are treated symmetrically. The two theories thus generate testable differences in trade patterns.²

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 role of contractible transfers, while Section 7 concludes.

2 The Environment

The economy consists of a principal and N agents, while time $t \in \{0, 1, \dots, \infty\}$ is discrete. At time 0 the principal designs a contract to maximise her expected profits. Each period

²Disclaimer: at the time of writing, Board (2006) has not yet been completed.

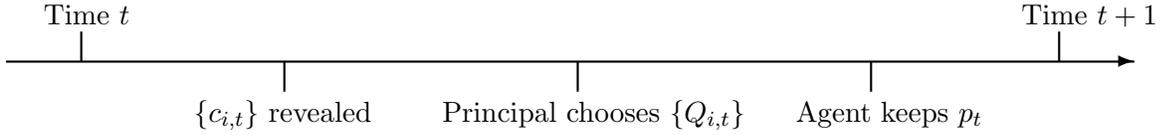


Figure 1: Timeline

$t \in \{1, \dots, \infty\}$ then consists of three stages. First, the cost of investing in agent i , $c_{i,t} \in [\underline{c}, \bar{c}]$, is publicly revealed. Second, the principal chooses to invest in one (or none) of the agents. This investment allows the agent to produce and sell a product worth v . Denote the probability that the principal invests in i by $Q_{i,t} \in \{0, 1\}$, where $\sum_i Q_{i,t} \leq 1$. Third, the winning agent chooses to keep p_t and pay the principal $v - p_t$, where the price p_t is noncontractible. The timeline is shown in Figure 1.

The stage game of this model exhibits the holdup problem. The agent has all the ex-post bargaining power, so they 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 make the natural assumption that contracts are maintained by bilateral reputations: that is, a deviation in the relationship between the principal and agent i cannot be observed by agents $j \neq i$.³ Formally, at time t , agent i observes $h_{i,t} := \{\{c_{i,t}\}_i, Q_{i,t}, p_t Q_{i,t}\}$. His 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 $Q_{i,t} : h^{t-1} \times [\underline{c}, \bar{c}]^N \rightarrow \{0, 1\}$, while the winning agent's pricing strategy is a mapping $p_t : h_i^{t-1} \times [\underline{c}, \bar{c}]^N \rightarrow [0, v]$.

A relational contract $\langle Q_{i,t}, p_t \rangle$ is defined to be a history-contingent plan of investments and prices. In Section 3 we assume the principal commits to her strategy, allowing us to focus on the agents' incentives. A contract is said to be *agent-self-enforcing* 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 a strategy. A contract is then said to be *self-enforcing* 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 from the principal's relationship with agent i is

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

³See Section 4 for discussion.

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

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

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

$$\Pi_{i,t} := E_t \left[\sum_{s=t}^{\infty} \delta^{t-s} (v - c_{i,t} - p_t) Q_{i,t} \right] \quad (2.1)$$

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

3 One-Sided Commitment

In this Section we suppose that the principal can commit to her investment strategy, $Q_{i,t}$. This acts as a useful initial benchmark and allows use to focus on the agents’ incentives.

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

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

Proof. First suppose that (DEA) holds. We use the following trigger punishment strategy: if i deviates at time t , then $Q_{i,s} = 0$ for $s > t$. If $Q_{i,t} = 1$, (DEA) implies that

$$E_t[\delta U_{i,t+1}] \geq v - p_t$$

so i will refrain from deviating. If $Q_{i,t} = 0$, the agent has no action and cannot deviate. Hence the contract is agent-self-enforcing.

Next suppose that a contract $\langle Q_{i,t}, p_t \rangle$ is agent-self-enforcing. The trigger punishment obtains the min-max of the stage game for agent i , so is an optimal penal code (Abreu (1988)). Subgame perfection then implies that there can be no one-shot deviation; that is, (DEA) must hold. \square

The *principal’s problem* is to choose the contract $\langle Q_{i,t}, p_t \rangle$ to maximise time-0 profit (2.1) subject to the agents’ dynamic enforcement constraint (DEA). This problem can be simplified by noting that the principal’s profit equals the welfare minus the sum of agents’ utilities. The principal’s problem is then to maximise

$$\Pi_0 := E_0 \left[\sum_{s=1}^{\infty} \sum_i \delta^{t-s} (v - c_{i,t}) Q_{i,t} \right] - \sum_i U_{i,0} \quad (3.1)$$

subject to (DEA).

Let $\tau_i(t) := \min\{s \geq t : Q_{i,s} = 1\}$ be the first time agent i trades after time t . 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*

$$U_{i,t} \geq E_t[v\delta^{\tau_i(t)-t}] \quad (\text{DEA}') \tag{3.1}$$

Proof. This 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 maximise profit (3.1) subject to (DEA'). The agents' utilities only enter profit via the time-0 term, so profit is maximised by ensuring that (DEA') binds at time-0,

$$U_{i,0} = E_0[v\delta^{\tau_i(0)}] \tag{3.2}$$

The *principal's problem* thus reduces to choosing $\langle Q_{i,t}, p_t \rangle$ to maximise profit

$$\Pi_0 := E_0 \left[\sum_{s=1}^{\infty} \sum_i \delta^{t-s} (v - c_{i,t}) Q_{i,t} \right] - E_0 \left[\sum_i v \delta^{\tau_i(0)} \right] \tag{3.3}$$

Notably, prices are absent from equation (3.3). This approach has allows us to solve the principal's problem in two steps. First, we choose the investment strategy $Q_{i,t}$ to maximise profits (3.3). Second, we use the agents' dynamic enforcement constraints (DEA) to derive prices. Since we have relaxed the principal's dynamic enforcement constraint, we call the solution to this problem the *optimal relaxed contract*.

3.1 The General Solution

The principal's problem (3.3) is an optimal-stopping problem and is hard to fully characterise. Fortunately, a number of economically relevant results immediately fall out.

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

Proposition 1. *Pick $i \in \mathcal{I}_t$. Suppose $c_{i,t} < v$, and*

$$c_{i,t} < c_{j,t} \quad \text{for all } j \in \mathcal{I}_t \tag{3.4a}$$

$$< c_{j,t} + (1 - \delta)v \quad \text{for all } j \notin \mathcal{I}_t \tag{3.4b}$$

*then $Q_{i,t} = 1$ in the optimal relaxed contract.*⁴

⁴Proposition 1 provides sufficient conditions for the principal to trade with agent i . One can also derive necessary conditions. Pick $i \in \mathcal{I}_t$. If either $c_{i,t} > v$ or $c_{i,t} > c_{j,t}$ for $j \in \mathcal{I}_t$, then $Q_{i,t} = 0$. Pick $i \notin \mathcal{I}_t$. If either $c_{i,t} > \delta v$ or $c_{i,t} > c_{j,t} + (1 - \delta)v$ for $j \in \mathcal{I}_t$, then $Q_{i,t} = 0$.

Proof. See Appendix A.1 □

The optimal relaxed contract has two key properties.

First, trade with insiders is efficient. Conditional on trading with an insider, the principal will trade with $i \in \mathcal{I}_t$ if and only if $c_{i,t} \leq \min\{c_{j,t}, v\}$ for all $j \in \mathcal{I}_t$. Intuitively, 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. 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.

Second, trade is biased against outsiders. If $i \notin \mathcal{I}_t$ and $c_{i,t} \leq \min\{c_{j,t}, v\}$ ($\forall j$), then the principal may not trade with i , even though it is efficient. Instead the principal may abstain from trading, go to an insider or even go to another outsider. Intuitively, the principal discriminates against outsiders because of the fixed cost of the relationship. If $(v - c_{i,t}) < (1 - \delta)v$ then the profit from agent i is less than the rental value of the rents the principal must pay him. Therefore, the principal prefers to abstain from trade rather than trade with agent i . Similarly, if $(c_{j,t} - c_{i,t}) < (1 - \delta)v$ for some $j \in \mathcal{I}_t$, then the principal will trade with insider, j , over the outsider because the cost differential is less than the rental value of i 's rents. Conditional on trading with an outsider, the principal may also prefer to trade with a relatively inefficient agent. Intuitively, the principal cares both about the cost in the current period and future trading opportunities, so she is biased towards agents who are more likely to trade sooner. Under IID costs, the future trading opportunity is the same for all agents, so the principal will prefer to use the most efficient outsider.

Proposition 1 provides a theory of endogenous switching costs. In typical models of switching costs (e.g. Klemperer (1995)), the principal must pay a fixed sum every time they change their current trading partner. In contrast, Proposition 1 says that the principal must pay a the sum when adopting a new partner, but not when reverting to an old partner.

It is also worth emphasising that the optimal contract is not stationary. In contrast, stationarity is optimal in relational contracting models with contractible transfers (e.g. MacLeod and Malcomson (1989)), and simply assumed in efficiency wage models (e.g. Shapiro and Stiglitz (1984), Greif (1993, 1994)).

3.2 Prices

The profit-maximising investment function $Q_{i,t}$ is chosen to maximise profits (3.3). Prices 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 (3.2). Given agents'

utilities, prices are then determined by backwards induction:

$$p_{i,t} = U_{i,t} - E_t[\delta^{\tau_i(t+1)} U_{i,\tau_i(t+1)}] \quad (3.5)$$

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 (3.5) the fastest prices are given by

$$p_{i,t} = v E_t[1 - \delta^{\tau_i(t+1)}] \quad (3.6)$$

A few features are worth noting. First, the time of the next trade is a sufficient statistic for the price. Second, prices increase as the time between trades grows. Third, 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.

Proposition 2. *Fix an investment 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

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

Utilities must obey (DEA'), so $\Pi_{i,t}$ is maximised by setting $U_{i,t} = E_t[v\delta^{\tau_i(t)-t}]$. □

Proposition 2 shows that the fastest prices maximise the principal's profit in each period. Consequently, when the principal cannot commit to the contract (see 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. The IID model is appealing because the identity of most efficient agent constantly changes, emphasising the tradeoff between maximising the gains from trade and maintaining a long run relationship. The IID case is also very tractable.

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 from insiders and from outsiders, denoted c^I and c^O respectively. For notational simplicity, assume that $\bar{c} \leq v$.

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, $c^I = \bar{c}$ and $c^O = \underline{c}$. Then it

must be the case that the principal prefers to use one last outsider (and paying rent v) rather than sticking to the current set of insiders. That is,

$$-\underline{c} + \frac{\delta}{1-\delta} E[v - c_{1:n^*}] \geq (v - \bar{c}) + \frac{\delta}{1-\delta} E[v - c_{1:n^*-1}] \quad (3.7)$$

Similarly, it must be the case that the principal does not want to increase the number of principals from n^* to $n^* + 1$, yielding another equality analogous to (3.7). Together these inequalities imply that n^* is the unique integer satisfying

$$\frac{\delta}{1-\delta} E[c_{1:n^*-1} - c_{1:n^*}] + (\bar{c} - \underline{c}) \geq v \geq \frac{\delta}{1-\delta} E[c_{1:n^*} - c_{1:n^*+1}] + (\bar{c} - \underline{c}) \quad (3.8)$$

Since the number of insiders can only increase over time, we have $n_t \rightarrow n^*$ almost surely. Two comparative statics are of interest. First, as agents become more patient, the benefit from future cost reductions increases and n^* rises. Second, as the value of trade v increases, so agents' rents increase and n^* declines.

The transition process can be characterised by backwards induction, using n^* as an initial condition. Denote the value function of the principals profit-maximisation problem in state n by $\Phi(n)$. This evolves as follows:

$$\begin{aligned} \Phi(n) &= E[v - c^I] + \delta\Phi(n) && \text{if } i \in \mathcal{I}_t \\ &= E[-c^O] + \delta\Phi(n+1) && \text{if } i \notin \mathcal{I}_t \end{aligned}$$

Denote the insider cost differential by $x := c^I - c^O$. The optimal policy then takes the form: choose the inside agent if the cost differential is less than some cutoff, $x < c_n^*$. The value function then becomes

$$\begin{aligned} \Phi(n) &= E[-c^O] + E[(v - x + \delta\Phi(n))\mathbf{1}_{x \leq c_n^*}] + E[\delta\Phi(n+1)\mathbf{1}_{x > c_n^*}] \\ &= \frac{E[-c^O] + E[(v - x)\mathbf{1}_{x \leq c_n^*}] + E[\delta\Phi(n+1)\mathbf{1}_{x > c_n^*}]}{1 - \delta E[\mathbf{1}_{x \leq c_n^*}]} \end{aligned} \quad (3.9)$$

The cutoff c_n^* is chosen to maximise $\Phi(n)$ and thus satisfies

$$(v - c_n^*) = \frac{\delta}{1-\delta} \left[(1-\delta)\Phi(n+1) - E[\max\{v - c_n^*, v - x\} - c^O] \right] \quad (3.10)$$

The left hand side of equation (3.10) 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 in the future, if $x > c_n^*$.

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 (3.10) to solve for the optimal state- $(n^* - 1)$ cutoff and equation (3.9) to calculate the value function $\Phi(n^* - 1)$. Then iterate.

Table 1 provides an illustration of this technique. 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, c_n^*	Prob($n_{t+1} = n_t$)	Value function, $\Phi(n)$
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, where $v = 2$, $c_{i,t} \sim [0, 1]$, $N = \infty$ and $\delta = 0.98$.

Throughout this Section we have assumed that the principal can commit to her strategy. This assumption is not ludicrous: in the example in Table 1, one can check that the principal has no incentive to renege on her investment plan. In Section 4 we investigate this issue in more detail.

4 No Commitment

In this Section we examine the infinitely repeated holdup game, where both the principal's and the agents' strategies form a subgame perfect equilibrium. We suppose that the principal's relationship with agent i is enforced by bilateral punishments. We use bilateral punishments because we wish to model a decentralised market where agent i 's information about agent j 's relationship is likely to be quite poor. This assumption has empirical relevance: McMillan and Woodruff's survey revealed that only 19% of Vietnamese firms thought that a cheating customer would be punished by other firms in the industry.

Lemma 3. *A contract $\langle Q_{i,t}, p_t \rangle$ is self-enforcing if and only if*

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

and

$$\Pi_{i,t} Q_{i,t} \geq 0 \quad (\forall i)(\forall t) \quad (\text{DEP})$$

Proof. Same as Lemma 1. □

The principal's problem is to maximise time-0 profit (2.1) subject to the dynamic enforcement constraints, (DEA) and (DEP). Appealing to Proposition 2, profits are maximised in each period if (DEA) binds ($\forall t$). The *principal's problem* thus reduces to choosing $\langle Q_{i,t}, p_t \rangle$ to maximise profit

$$\Pi_0 := E_0 \left[\sum_{s=1}^{\infty} \sum_i \delta^{t-s} (v - c_{i,t}) Q_{i,t} \right] - E_0 \left[\sum_i v \delta^{\tau_i(0)} \right] \quad (4.1)$$

subject to

$$(W_{i,t} - v) Q_{i,t} \geq 0 \quad (\forall i)(\forall t) \quad (\text{DEP}')$$

Notably, prices are absent from both (4.1) and (DEP'). We call the solution to this problem the *optimal contract*.

4.1 Time Inconsistency

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 costs $c_{1,t} \in (0, v)$ are deterministic and increasing such that $\lim_{t \rightarrow \infty} c_{1,t} = v$ and $\sum_{t=1}^{\infty} \delta^t (v - c_{1,t}) > v$. In this case, the optimal relaxed contract is $Q_{i,t} = 1$ ($\forall t$). However, since the gains from trade disappear over time, the only solution to satisfy (DEP) is $Q_{1,t} = 0$ ($\forall t$). This unravelling result holds for all $\delta \in (0, 1)$. △

The problem is that the principal's optimal relaxed 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 will 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.

4.2 IID Costs with Patient Agents

To further examine the principal's optimal contract let us return to the IID benchmark. Proposition 3 shows that, when the principal is sufficiently patient, the optimal contract is self-enforcing.

Proposition 3. *Suppose that costs are IID. In addition, assume that either $\underline{c} > 0$ or $v > \bar{c}$. Then there exists a $\hat{\delta} < 1$, independent of N , such that the optimal relaxed contract satisfies (DEP) when $\delta > \hat{\delta}$.*

Proof. See Appendix A.2. □

Example 2 (Uniform Costs). Suppose, $N = \infty$, $c_{i,t} \sim U[0, 1]$ and $v > 1$. As shown in Appendix A.3, (DEP) is satisfied if $\delta \geq \hat{\delta} := 1/(1 + (v - 1)^3)$. For example, if $v = 2$, then $\hat{\delta} = 1/2$. As $v \rightarrow 1$, so $\hat{\delta} \rightarrow 1$. △

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

Proposition 3 makes a stronger statement: there exists a critical discount factor, $\hat{\delta}$, independent of the number of agents, such that the optimal contract satisfies (DEP) when $\delta \geq \hat{\delta}$. When there are an infinite number of agents, the problem is that the number of trading partners will increase without bound as $\delta \rightarrow 1$. 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 (3.7). 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.

4.3 IID Costs with Impatient Agents

Proposition 3 shows that, when agents are patient, the principal can implement the optimal relaxed contract. However, when agents are impatient, efficient trade may not satisfy the principal's dynamic enforcement constraint. The principal can respond by introducing trading inefficiencies in a number of ways. Proposition 4 illustrates how inefficiency can ease the dynamic enforcement constraint in the case when $N = 1$.⁵

⁵With additional agents, the problem is much harder. This is the subject of the companion paper, Board (2006).

First, recall that the optimal relaxed contract is as follows:⁶

$$\begin{aligned} Q_t &= \mathbf{1}_{c_t \leq c^*} && \text{if } i \notin \mathcal{I}_t \\ Q_t &= \mathbf{1}_{c_t \leq v} && \text{if } i \in \mathcal{I}_t \end{aligned}$$

where $c^* \geq \delta v$.

Proposition 4. *Suppose costs are IID and $N = 1$. Then the optimal contract obeys*

$$\begin{aligned} Q_t &= \mathbf{1}_{c_t \leq \kappa^*} && \text{if } i \notin \mathcal{I}_t \\ Q_t &= \mathbf{1}_{c_t \leq \kappa^{**}} && \text{if } i \in \mathcal{I}_t \end{aligned}$$

where $\kappa^* \leq \kappa^{**}$. In addition, $\kappa^* \leq c^*$ and $\kappa^{**} \leq v$.

Proof. See Appendix A.4. □

When the principal cannot commit, the time inconsistency problem implies that the principal will wish to renege on a trade if the future value of the relationship is less than the cost, c_t . Credibility can thus be enhanced through a contract where trade does not occur when c_t exceeds some cutoff, κ^{**} . Since efficient trading opportunities are sometimes forgone, the ex-ante value of trade falls, reducing the willingness of the principal to invest in a new agent, and hence reducing κ^* .

5 Private Cost Information

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

5.1 Failure of Incentive Compatibility

In Section 4 we showed that, if agents are sufficiently patient, the optimal relaxed contract is self-enforcing. It will not, however, tend to be incentive compatible.⁷

To see the problem suppose that the principal promises to implement the optimal relaxed contract using some prices $\{p_t\}$. The principal may have two reasons to distort her cost reports:

⁶Since there is only one agent, we drop i subscripts.

⁷In the text, our presentation is heuristic. To be more formal, consider the following game. First, the principal privately observes her costs. Second, she makes a public cost report which determines the contract $(Q_{i,t}, p_t)$. Third, she chooses whether or not to invest according to the plan. Fourth, the winning agent chooses whether or not to hold the principal up. A contract is incentive compatible if the principal cannot improve her payoffs by misreporting at stage 2.

1. Suppose $p_t > v - c_{i,t} > 0$ for $i \in \mathcal{I}_t$, and $Q_{i,t} = 1$. The principal may then prefer to exaggerate $c_{i,t}$ to avoid trade with agent i .
2. Suppose $c_{i,t} > c_{j,t}$ for $i \in \mathcal{I}_t$ and $j \notin \mathcal{I}_t$, and $Q_{i,t} = 1$. The principal may then prefer to exaggerate $c_{i,t}$ or under-report $c_{j,t}$ in order to trade with j rather than i .

These problems result from the time-inconsistency of the optimal relaxed contract, under which the principal treats insiders efficiently. 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 cost is 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, or restrict the agents with whom the principal could trade. 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 relaxed contract.

5.2 Maintenance Payments

So far we have assumed that transfers between the principal and agent i can only be made when $Q_{i,t} = 1$. Instead suppose that transfers can take place in any period after the relationship has been initiated, i.e. after $\tau_i(0)$.⁸ Denote the transfer paid in period t to agent i by $p_{i,t}$.

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

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

The investment strategy $Q_{i,t}$ is then chosen to maximise time-0 profits (3.3).

Proposition 5. *Suppose that costs are privately known by the principal. The maintenance contract (5.1) is an optimal relaxed contract and is incentive compatible. Moreover, it is self-enforcing if*

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

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

⁸In Section 6 we argue that transfers prior to $\tau_i(0)$ are unlikely to occur. These criticisms do not apply to transfers after $\tau_i(0)$.

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

$$\Pi_t = E_t \left[\sum_{s=t}^{\infty} \delta^{s-t} (v - c_{I_t,t}) \right] - E_t \left[\sum_{i \notin \mathcal{I}_t} \delta^{\tau_i(t)-t} v \right] - \sum_{i \in \mathcal{I}_t} v \quad (5.2)$$

Observe that the last term in (5.2) is sunk and can be ignored. The investment plan, $Q_{i,t}$, chosen to maximise time-0 profits (3.3) thus maximises time- t profits (5.2) 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,

$$\Pi_{i,t} = W_{i,t} - v = E_t[\delta W_{i,t+1}] - c_{i,t} \geq 0 \quad (5.3)$$

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,

$$\Pi_{i,t} = W_{i,t} - v = E_t[\delta W_{i,t+1}] - v \geq 0 \quad (5.4)$$

When $Q_{i,t} = 0$ and $i \notin \mathcal{I}_t$, the principal cannot defect. Putting (5.3) and (5.4) together, yields (DEP^{MC}) \square

The maintenance contract pays each insider $(1 - \delta)v$ per period, independent of whether trade occurs or not. Since these maintenance payments are sunk, the principal's problem is time consistent. As a consequence, the principal has no incentive to lie about her costs.

Proposition 5 is very general. We assumed nothing about the stochastic structure of costs. The mechanism 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 maintenance payments is that the principal's dynamic enforcement constraint (DEP^{MC}) is stricter than under the fastest price contract (DEP). The reason for the difference is straightforward: under the maintenance payment contract, insiders receive utility $U_{i,t} = v$ whether or not they trade; under the fastest price contract, insiders receive $U_{i,t} = v$ when they trade and receive strictly less in other periods. Nevertheless, when costs are IID, Proposition 6 shows that the difference between (DEP) and (DEP^{MC}) is relatively minor.

Proposition 6. *Suppose that costs are IID. In addition, assume that either $\underline{c} > 0$ or $v > \bar{c}$. Then there exists a $\hat{\delta} < 1$, independent of N , such that the optimal relaxed contract satisfies (DEP^{MC}) when $\delta > \hat{\delta}$.*

Proof. Same as proof of Proposition 3. \square

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

6 On Transfers

This paper has argued that the principal will allocate investment in order to minimise the rents obtained by agents. However, the reader has probably observed that, by defining the space of payments more broadly, the principal can extract all rents from the agents. There are two ways this can be achieved:

1. Allow the agents to make contractible payments to the principal. By setting this contractible payment equal to v , the principal can then extract all rents from the agents, as shown by MacLeod and Malcomson (1989).
2. Allow the agent to make a voluntary payment to the principal before time $\tau_i(0)$. By setting this voluntary payment equal to the expected value of the agents future rents, the principal can again fully extract from the agents.

There are two basic problems with contracts involving full-extraction transfers of this sort. The first problem is practical. In many developing countries, such as Vietnam, contractible payments cannot be implemented since a there is no fully functioning legal system. While some firms make up-front payments, they are far from universal. In their survey of Eastern European firms, Johnson, McMillan, and Woodruff (2002) find that only 35% ask for any advance payment. When looking at custom built products, only 42% ask for any advance payment.

A second problem is more theoretical: these full-extraction transfers are not robust to free entry into the market for principals. Suppose a contract $\langle Q_{i,t}, \phi_{i,t}, \phi_{i,t}^0 \rangle$ consists of three components: the principal's investment decision, $Q_{i,t}$; a voluntary payment from agent i to the principal, $\phi_{i,t}$; and a contractible payment from agent i to the principle, $\phi_{i,t}^0$. We say a contract is *cowboy-proof* if a new principal, with costs $c_{i,t} = \infty$, cannot make a strictly positive profit.⁹

Proposition 7. *Profits under the optimal self-enforcing cowboy-proof contract $\langle Q_{i,t}, \phi_{i,t}, \phi_{i,t}^0 \rangle$ equal profits under the optimal contract $\langle Q_{i,t}, p_t \rangle$.*

⁹This argument is quite intuitive. After moving across the country to accept a new job, buying a house and enrolling their children in the local schools, a professor can be held up by their university. While the professor could ask for a large up-front payment to compensate for the low wages, a university will be very hesitant about making such a payment. In particular, they will fear that the professor will accept the up-front payment and leave the following year. This argument is analogous to the free entry model of Shapiro (1983)

Proof. Recall that the optimal contract $\langle Q_{i,t}, p_t \rangle$ maximises profit (4.1) subject to (DEP'). Since all payments are voluntary, the optimal contract is thus cowboy-proof.

Let us turn to the optimal self-enforcing cowboy-proof contract $\langle Q_{i,t}, \phi_{i,t}, \phi_{i,t}^0 \rangle$. Cowboy-proofness says that a cowboy's profits from agent i ,

$$\Pi_{i,0}^C := E_0 \left[\sum_{t < \tau_i(0)} \delta^t (\phi_{i,t} + \phi_{i,t}^0) + \delta^{\tau_i(0)} \phi_{i,t}^0 \right]$$

must be negative. By holding-up a principal who invests at time $\tau_i(0)$, an agent can then obtain rents

$$U_{i,0} \geq E_0[v\delta^{\tau_i(0)}] - \Pi_{i,0}^C \geq E_0[v\delta^{\tau_i(0)}]$$

Summing over i , profits are thus weakly larger than those in (4.1).

Let us now examine the dynamic enforcement constraints. Pick any time t and agent i such that $Q_{i,t} = 1$. The agent will not defect if

$$U_{i,t} \geq v - \phi_{i,t}^0 \tag{6.1}$$

Similarly, the principal will not defect if

$$\Pi_{i,t} \geq \phi_{i,t}^0 \tag{6.2}$$

Summing (6.1) and (6.2) yields (DEP'). Putting this together, the optimal contract $\langle Q_{i,t}, p_t \rangle$ attains the maximum profit from any self-enforcing cowboy-proof contract. \square

Proposition 7 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 (a) payments are voluntary, and (b) payments are only made when trade takes place. Thus expanding the space of possible payments does not benefit the principal.¹⁰

7 Conclusion

In countries with poor contract law, relational contracts are essential to encourage investment and facilitate trade. While relational contracts perform far better than spot markets, this paper has sought to examine how these ongoing relationships can create endogenous switching costs and distort the pattern of trade in the economy.

The theory we have put forward is, in essence, very simple. Agents' ability to holdup the

¹⁰This, of course, ignores issues of incentive compatibility. It is worth noting that the maintenance contract is also cowboy-proof.

principal enables the agents to extract rents. Since an agent cares about their utility from the entire trading relationship, the same rent can also be used to prevent multiple incidences of holdup. That is, the rent acts like a fixed cost of initiating a new relationship. In response to this fixed cost, the principal starts fewer new relationships than is first-best.

We also showed that the optimal contract is time inconsistent, giving rise to two related problems. First, the principal may wish to renege on the optimal contract. However, this problem can be overcome if the principal is sufficiently patient, no matter how many agents there are in the industry. Second, the principal may wish to lie about her gains from trade in order to undo the optimal contract. However, this problem can be overcome if the principal can use a maintenance contract.

Our theory has testable implications. First, firms should be more willing to revert to an old supplier than use a new supplier. Second, a firm may refuse to exploit new trading opportunities if they must source from a new partner. Third, switching costs should decrease with agents' patience, but should never disappear. Our theory also helps compare the performance of relational contracting and an efficient legal system. It suggests, for example, that formal contracts are most important where the efficient trading partner changes frequently, and where there are many potential partners.

A Omitted Proofs

A.1 Proof of Proposition 1

Fix t and suppose that $i \in \mathcal{I}_t$ obeys (3.4a)–(3.4b). By the principal of optimality, the principal wishes to maximise

$$\Pi_t = E_t \left[\sum_{s=t}^{\infty} \sum_i \delta^{s-t} (v - c_{i,s}) Q_{i,s} \right] - E_0 \left[\sum_{k \notin \mathcal{I}_t} \delta^{\tau_k(t)} v \right] \quad (\text{A.1})$$

Since $c_{i,t} < v$, the principal prefers trading to abstaining. By (3.4a) the principal also prefers trading with i to trading with a different $j \in \mathcal{I}_t$. To show the principal prefers to trade with i than $j \notin \mathcal{I}_t$, denote the profit from trading with i by $\Pi_t(i)$. We then have,

$$\begin{aligned} \Pi_t(i) &= (v - c_{i,t}) + \delta E_t \left[\sum_{s=t+1}^{\infty} \sum_i \delta^{s-t} (v - c_{i,s}) Q_{i,s} \right] - E_0 \left[\sum_{k \notin \mathcal{I}_t} \delta^{\tau_k(t+1)} v \right] \quad (\text{A.2}) \\ &\geq (v - c_{i,t}) + \delta E_t \left[\sum_{s=t+1}^{\infty} \sum_i \delta^{s-t} (v - c_{i,s}) Q_{i,s} \right] - \delta v - E_0 \left[\sum_{k \notin \mathcal{I}_t \setminus \{j\}} \delta^{\tau_k(t+1)} v \right] \\ &\geq (v - c_{j,t}) + \delta E_t \left[\sum_{s=t+1}^{\infty} \sum_i \delta^{s-t} (v - c_{i,s}) Q_{i,s} \right] - v - E_0 \left[\sum_{k \notin \mathcal{I}_t \setminus \{j\}} \delta^{\tau_k(t+1)} v \right] = \Pi_t(j) \end{aligned}$$

where the third line follows from (3.4b).

A.2 Proof of Proposition 3

We seek to show that, for δ sufficiently high, (DEP) is slack under the optimal contract. If $Q_{i,t} = 1$, the profit from relationship i is

$$\begin{aligned} \Pi_{i,t} &= -c_{i,t} + \delta E_t[W_{i,t+1}] \\ &\geq -v + \frac{1}{n^*} \frac{\delta}{1 - \delta} \omega(n^*) \end{aligned} \quad (\text{A.3})$$

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 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. \square

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

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

$$n_i \Delta\omega(n_i) \geq k - \epsilon > 0 \quad (\forall i) \tag{A.4}$$

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.4), while the fifth line follows from Lemma 4. \square

Using the same approach as behind equation (3.7), we can characterise the maximum number

of agents, n^* .¹¹

$$\frac{\delta}{1-\delta} \Delta\omega(n^*) \geq \max\{v - \bar{c}, 0\} + \underline{c} \quad (\text{A.5})$$

Substituting (A.5) into (A.3),

$$\Pi_{i,t} \geq -v + \frac{\omega(n^*)}{n^* \Delta\omega(n^*)} [\max\{v - \bar{c}, 0\} + \underline{c}] \quad (\text{A.6})$$

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

A.3 Derivation of Example 2

Suppose $c_{i,t} \sim U[0, 1]$ and $v > 1$. As in equation (A.3), profits at time t are

$$\begin{aligned} \Pi_{i,t} &= -c_{i,t} + \delta E_t[W_{i,t+1}] \\ &\geq -1 + \frac{1}{n^*} \frac{\delta}{1-\delta} \omega(n^*) \\ &\geq -1 + \left[\frac{1}{n^*(n^*+1)} \right]^{1/2} \frac{\delta}{1-\delta} (v-1) \end{aligned} \quad (\text{A.7})$$

since $\omega(n) = v - \frac{1}{1+n}$. Using equation (A.5),

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

Rearranging and substituting into (A.7) 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.4 Proof of Proposition 4

The problem is to maximise (4.1) subject to (DEP'). First, suppose that the agent is an insider (i.e. that they have bought previously). By the principal of optimality, the principal wishes to

¹¹Of course, this bound will not be reached if $n^* > N$.

maximise

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

subject to

$$(\delta E_t[W_{t+1}] - c_t)Q_t \geq 0 \quad (\text{DEP}')$$

Since this is a stationary problem, the solution will be stationary. Moreover, the solution will be monotone since trading when the cost is low raises the objective and weakens the constraint. We can thus define $W(\kappa)$ by

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

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

$$\kappa^{**} = \max\{\kappa \leq v : \delta E[W(\kappa)] - \kappa \geq 0\} \quad (\text{A.8})$$

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

$$\begin{aligned} \Pi(\kappa^*) &= \int_{\underline{c}}^{\kappa^*} (\delta E[W(\kappa^{**})] - c) dF(c) + \int_{\kappa^*}^{\bar{c}} \delta \Pi(\kappa^*) dF(c) \\ &= \frac{\int_{\underline{c}}^{\kappa^*} (\delta E[W(\kappa^{**})] - c) dF(c)}{1 - \delta(1 - F(\kappa^*))} \end{aligned} \quad (\text{A.9})$$

The initial cutoff κ^* is chosen to maximise $\Pi(\kappa^*)$ subject to (DEP'). Since (DEP') must hold, $\kappa^* \leq \kappa^{**}$. Finally, we wish to show that $\kappa^* \leq c^*$ and $\kappa^{**} \leq v$. The latter follows from (A.8). 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 relaxed contract, the first period cutoff is higher, as required.

References

- ABREU, D. (1988): “On the Theory of Infinitely Repeated Games with Discounting,” *Econometrica*, 56(2), 383–396.
- BOARD, S. (2006): “Reputation and Loyalty with Impatient Agents,” Working paper, University of Toronto.
- GREIF, A. (1993): “Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders’ Coalition,” *American Economic Review*, 83(3), 525–548.
- (1994): “Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies,” *Journal of Political Economy*, 102(5), 912–950.
- 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.
- JOHNSON, S., J. McMILLAN, AND C. WOODRUFF (2002): “Courts and Relational Contracts,” *Journal of Law, Economics and Organization*, 18, 221–277.
- KLEIN, B., AND K. B. LEFFLER (1981): “The Role of Market Forces in Assuring Contractual Performance,” *Journal of Political Economy*, 89(4), 615–641.
- 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.
- 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.
- 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.
- McMILLAN, J., AND C. WOODRUFF (1999a): “Dispute Prevention without Courts in Vietnam,” *Journal of Law, Economics and Organization*, 15(3), 637–658.
- (1999b): “Interfirm Relationships and Informal Credit in Vietnam,” *Quarterly Journal of Economics*, 114(4), 1285–1320.

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.