I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15881

Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment: Thirty Years On

W. Bentley MacLeod
James M. Malcomson

JANUARY 2023

IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15881

Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment: Thirty Years On

W. Bentley MacLeod
Princeton University, NBER and IZA

James M. Malcomson
University of Oxford and IZA

JANUARY 2023

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment: Thirty Years On

"Implicit Contracts, incentive compatibility, and involuntary unemployment" (MacLeod and Malcomson, 1989) remains our most highly cited work. We briefly review the development of this paper and of our subsequent related work, and conclude with reflections on the future of relational contract theory and practice.

JEL Classification: D21, D23, D82, D86, L14, L22, L23, L24

Keywords: relational contracts, informal enforcement, legal enforcement, incentives, private information

Corresponding author:
W. Bentley MacLeod
Princeton University
Princeton
New Jersey 08544
USA

E-mail: wbmacleod@wbmacleod.net

"Implicit Contracts, incentive compatibility, and involuntary unemployment" (MacLeod and Malcomson, 1989) remains for both of us our most highly cited work. In this paper we briefly review the development of that paper and its key contributions, along with a brief discussion of our subsequent related work. We finally reflect on current trends in the field and what we see as important open questions.

1 Intellectual Origins

MacLeod and Malcomson (1989) brings together some ideas that we had been working on separately. The first of these is the observation that labor market contracts are incomplete. Malcomson (1981) makes this explicit. It is worthwhile quoting exactly what is said in his 1981 paper:

> In fact many, if not most, employment contracts between a firm and its employees do not specify the productivity level at all precisely. Contracts based on payment by time rate may specify certain hours of work but rarely specify precisely everything that is to be done during those hours. Contracts based on payment by piece rate rarely specify the rate at which output is to be produced, which is not in general a matter of indifference to the firm. One reason that has been given for this incompleteness of contracts is what Simon (1979) and Williamson et al. (1975) have called bounded rationality, which in this context can be interpreted as the limited capacity of human beings to conceive of, let alone write contracts to cover, every contingency that may be relevant. Loasby (1976) argues that this results in employment contracts that are generally vague as to the precise duties involved.

The idea that contracts are incomplete applies not only to employment but also to supply and service contracts, as documented by Macaulay (1963). Malcomson (1984, 1986) explored relative performance evaluation (tournaments) as a way of handling such incompleteness. Malcomson (1981), however, concluded by observing that a natural implication of incomplete contracts is involuntary unemployment. The subsequent paper by Shapiro and Stiglitz (1984) provides a more elegant version of this idea that in the end became the standard model of efficiency wages in the labor economics literature. In some sense these ideas were not new; as Bowles (1985) observes, they go back to Marxian notions of a reserve army. What the formal model provides is a framework in which we can see how the various economic forces affect important observables, such as the wage rate and the level of unemployment. In particular, the efficiency wage model highlighted the fact that individuals can be motivated to work based upon the future expectation of employment at a high wage.

There were, however, unresolved issues with the Shapiro and Stiglitz (1984) model. Carmichael (1985) observes that logically, the efficiency wage model does not predict involuntary unemployment because firms can find methods, such as the discretionary bonuses analyzed by Bull (1987), to ensure that the expected utility of workers is equal to their next best alternative. What is missing here is a more comprehensive view of the interaction between markets and the problems that arise due to incomplete contracts. Some insight into this problem is provided by Hirschman (1970)'s ideas on exit and voice, later reflected in the work of Freeman and Medoff (1984).

Hirschman's discussion is based on the observation that organizations such as state schools rely on the active participation of individuals in the community, an effect he calls *voice*. The problem is that if there is an increase in competition from other jurisdictions, individuals may leave the community, which can send quality into a downward spiral. Bentley attended a lecture by Jan Svenjar on labor managed firms (see Svenjar (1982)), and wondered how they survived the problem of exit. As it turns out, further research revealed that successful cooperatives typically have some form of exit cost. His earlier work on cooperation between firms (MacLeod (1985)) led to the idea to model cooperation as a repeated game and to show that a necessary condition for successful cooperation is some form of exit cost (MacLeod (1984), later published as MacLeod (1987) and MacLeod (1988)). In the Shapiro and Stiglitz (1984) model, involuntary unemployment acts as a cost of exiting an employment relationship.

Jim and Bentley met at CORE in Belgium in the early 1980s and discussed putting these ideas together. The final ingredients were the work of Abreu (1988), who built on Friedman (1971), and ideas from dynamic programming to provide a complete characterization of the set of equilibria in a repeated game (see Thomas and Worrall (2022)'s discussion on the impact of dynamic programming for relational contract theory). We combined this idea with a model of incomplete contracts to provide a complete characterization of the set of self-enforcing agreements in a competitive market. The next section provides a brief review of the ideas.[1]

[1] For comprehensive reviews of the relational contracts literature applied to labor markets and supply relationships, see Malcomson (1999) and Malcomson (2013) respectively. See MacLeod (2007) for a discussion of the enforcement of incomplete contracts under different information conditions, and the consequences for

2 A Basic Model

In the simplest terms, the fundamental contributions of MacLeod and Malcomson (1989) were: (1) to show that the costs of reneging on an agreement (or, equivalently, the economic gains or rents from continuing it), known from the repeated game literature to be a requirement for sustaining a cooperative outcome, affect what is sustainable in a particularly simple way in a relational contract with money transfers between the parties; (2) to characterize the minimum reneging costs required to sustain an outcome; and (3) to fully characterize when market equilibria can sustain those costs. Importantly in a labor market context, whenever an agreement can be sustained, there are market equilibria with just involuntary unemployment (the employee faces all the reneging costs), market equilibria with involuntarily unfilled vacancies but no involuntary unemployment (the employer faces all the reneging costs), and market equilibria with some of each.

A somewhat more formal understanding of these results is provided by a model with a purchaser (employer, downstream firm) buying from a supplier (employee, upstream firm) when there is no measure of the supplier's performance that is verifiable in court. (Readers who want to avoid technicalities should skip to Theorem 1 below.) The purchaser's profit in period t from supplier performance $q_t \in [0,\overline{q}]$ is $q_t - P_t$, where P_t is the payment to the supplier. Performance can be multi-dimensional, with q_t a vector. It is essentially a shorthand term for anything under the control of the supplier that is of concern to the purchaser. The supplier incurs cost $c_t(q_t)$ for performance q_t and so receives payoff $P_t - c_t(q_t)$ in period t if paid P_t for delivering q_t. Conventionally, c_t is continuous, strictly increasing and strictly convex (so, whenever $c_t(q_t)$ is twice differentiable, $c_t'(q_t) > 0$ and $c_t''(q_t) > 0$ for all q_t), $c_t(0) = 0$, and $\lim_{q_t \to \overline{q}} c_t(q_t) = \overline{c}$ with $\overline{c}$ finite. If the parties choose not to trade with each other in period t, the purchaser receives payoff $\pi_t > 0$ (perhaps from producing the component in-house) and the supplier payoff $u_t > 0$ (perhaps from supplying elsewhere). The parties have a common discount factor $\delta_t \in (0,1)$.

When performance is unverifiable, it is not possible to write a legally enforceable contract that makes payment conditional on it. However, as Bull (1987) and MacLeod and Malcomson (1989) observed, the parties can still use payments conditional on performance provided the relationship is structured in such a way that it is worthwhile for the purchaser to make them. For that reason, it may be useful to have the payment P_t consist of two components, a legally enforceable component p_t that cannot be conditioned on performance

the interaction between enforcement costs and the efficiency of incomplete contracts.

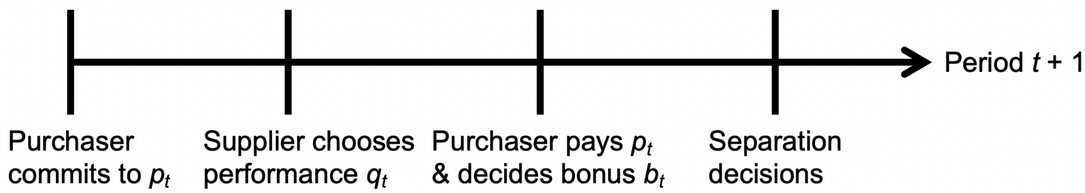

Figure 1: Timing of events for period t

and a discretionary bonus $b_t \geq 0$ that can. Following Levin (2003), we can think of the purchaser and supplier agreeing informally a plan for q_t, p_t and b_t for the duration of their relationship. This is their *relational contract*.

The timing of events for period t is shown in Figure 1. At the beginning of each period, the purchaser commits to the legally enforceable payment p_t.[2] The supplier then decides performance. The purchaser observes performance, pays p_t and decides whether to pay b_t. Finally in this period, the relationship may be terminated by either party.

With this structure, no supply occurs if the relationship lasts for at most one period. With no future to be concerned about, the purchaser has no incentive to pay a positive bonus. Anticipating this, the supplier has no incentive to deliver performance above zero. With $q_t = 0$, the combined payoffs from the relationship are zero, which is less than the combined payoffs $\pi_t + u_t$ the parties could get by not trading with each other, so there is no payment for which they would trade. Thus the possibility of trade depends on the relationship being at least potentially on-going.

For an on-going relationship, let Π_t and U_t respectively denote the purchaser's and the supplier's future payoffs from t on if both parties stick to their relational contract, $\tilde{\Pi}_t$ and $\tilde{U}_t$ their worst future payoffs if they cheat on it, and $\bar{\Pi}_t$ and $\bar{U}_t$ their payoffs if the relationship ends for reasons unrelated to the performance of either party. (In principle, $\bar{\Pi}_t$ and $\bar{U}_t$ need not be the same as $\tilde{\Pi}_t$ and $\tilde{U}_t$.) The differences $\Pi_t - \tilde{\Pi}_t$ and $U_t - \tilde{U}$ are the gains to purchaser and supplier respectively of not cheating on the relational contract. A central result that follows from MacLeod and Malcomson (1989) is:

Theorem 1. *(MacLeod and Malcomson, 1989) Necessary and sufficient conditions for a self-enforcing relational contract with payoffs Π_t and U_t with performance q_t are:*

Incentive compatibility: *For all t, the discounted joint future payoff gain to purchaser and supplier from continuing the relational contract is at least as great as the cost of*

[2]The need for p_t to be legally enforceable could be eliminated by having it paid at this stage.

supplying performance q_t:

$$\delta_t \left(U_{t+1} - \tilde{U}_{t+1} + \Pi_{t+1} - \tilde{\Pi}_{t+1} \right) \geq c_t \left(q_t \right), \quad \textit{for all } t. \tag{1}$$

Individual rationality: *For all t, the payoffs to purchaser and supplier from continuing the relational contract are at least as great as their alternatives:*

$$\Pi_t \geq \bar{\Pi}_t, \quad U_t \geq \bar{U}_t, \quad \textit{for all } t. \tag{2}$$

For a sketch of the proof, see Appendix A. MacLeod and Malcomson (1989) then show that market equilibrium places no restrictions on the allocations beyond those required by Theorem 1 with, for all t, $\tilde{\Pi}_t = \bar{\Pi}_t$ the payoff to the purchaser from leaving the market altogether and $\tilde{U}_t = \bar{U}_t$ the payoff to the supplier from leaving the market altogether.

The right-hand side of (1) is the cost of delivering performance q_t. That is sometimes referred to as the *reneging temptation*. The left-hand side of (1) is the sum of the discounted future gains to the parties from continuing the relationship, the *reneging costs*, at the end of period t. Thus the incentive conditions to sustain performance q_t take the simple form that only the *sum* of the reneging costs, not which party would incur them, exceeds the reneging temptation. The right-hand side of (1) specifies the minimum value of that sum. Moreover, the sum of the future gains is independent of the payments made by the purchaser to the supplier — these cancel when U_{t+1} and Π_{t+1} are added together — so the only element of the relational contract that affects whether (1) is satisfied is the performance sequence $(q_t, q_{t+1}, \ldots)$ from t on. For any performance sequence that satisfies (1), it is always possible to find a bonus sequence $(b_t, b_{t+1}, \ldots)$ that distributes those gains in any way between purchaser and supplier, so there are multiple equilibria. Only the joint future gain, not its division between purchaser and supplier, enters (1) for the following reason. Future gain to the supplier affects the supplier's incentive for performance. But future gain to the purchaser can do so too because it can be used to induce the purchaser to pay a bonus and that bonus affects the performance that can be delivered. When reneging results in the relationship ending (so $\tilde{U}_t$ and $\tilde{\Pi}_t$ correspond to market alternatives), at least one party must be strictly better off having the relationship continue than from what they can obtain elsewhere. In an employment context with the number of matches maximized, that corresponds to either involuntary unemployment or involuntarily unfilled vacancies. Specifying a severance payment from one party to the other only redistributes gains; it does not affect sustainable performance.

As noted above, in a one-shot game the purchaser would never pay a positive bonus, so the supplier would always set performance at the lowest level. Thus, the one-shot game has a prisoners' dilemma type structure and one can apply the intuition for the standard Folk Theorem result in a repeated prisoners dilemma game. Suppose δ_t is constant at δ for all t. The Folk Theorem result tells us that, for a repeated prisoners dilemma game, any individually rational cooperation can be sustained for a discount factor δ sufficiently close to 1 by strategies that revert to those of the one-shot equilibrium if either party deviates. Consider any performance sequence $(q_t, q_{t+1}, \ldots)$ from t on for which the left-hand side of (1) is positive for all t. As $\delta \to 1$, that left-hand side goes to infinity so any such sequence satisfies (1). Equality in (1) specifies the critical value of the discount factor for sustaining any particular sequence.[3] Provided this performance sequence is mutually beneficial, there also exists a sequence $(p_t, p_{t+1}, \ldots)$ such that individual rationality for both parties (2) is satisfied for all t. Then the on-going nature of the relationship enables the purchaser to elicit performance from the supplier above the minimum level when that would not be possible in a one-period relationship.

3 Implications and Subsequent Joint Research

3.1 Payoffs on the Pareto Frontier

An implication of Theorem 1 is that any performance sequences that satisfy (1) and (2) can be sustained as equilibria. But some equilibria may be better than others. Let q_t^* denote first-best performance defined by

$$q_t^* = \arg\max_q q - c_t(q), \quad \text{for all } t \geq \tau. \tag{3}$$

The purchaser and supplier can increase their joint payoff $U_\tau + \Pi_\tau$ at the start of a relationship at τ unless, for each $t \geq \tau$, either (a) $q_t < q_t^*$ and (1) holds with equality, or (b) $q_t = q_t^*$. To see why, note that the component of the joint payoff at date t', $U_{t'} + \Pi_{t'}$, arising at date $t \geq t'$ is just $q_t - c_t(q_t)$, which q_t^* maximizes. Thus the joint payoff $U_{t'} + \Pi_{t'}$ is always increased by having q_t move towards q_t^* for $t \geq t'$. So moving q_t towards q_t^* increases the

[3]But see Blonski et al. (2011) on whether this value is actually critical for determining whether the parties sustain such a sequence. See also Abreu et al. (1991) who show that this result does not necessarily hold with imperfect information. MacLeod (2007) discusses the implications of this result for relational contract theory.

left-hand side of (1) for all dates prior to t. Suppose $q_t > q_t^*$. Then the parties can always increase their joint payoff at τ because reducing q_t increases the joint payoff for all dates up to t, so relaxing the constraint (1) for all dates prior to t by increasing the left-hand side, and also relaxes the constraint (1) for t by reducing the right-hand side. Appropriate choice of p_t and b_t can ensure that neither U_t nor Π_t is reduced.

Suppose now $q_t < q_t^*$. If (1) held with strict inequality at t, q_t could be increased without violating that constraint and this would also increase the left-hand side of (1) for all dates prior to t, so relaxing the constraint for all dates prior to t. Thus increasing q_t would increase the joint payoffs at all dates up to t without resulting in any of the constraints (1) and (2) being violated.

This argument establishes that any self-enforcing equilibrium on the Pareto frontier of payoffs has, for each $t \geq \tau$, either (a) $q_t < q_t^*$ and (1) holding with equality, or (b) $q_t = q_t^*$. This and pooling of incentive compatibility conditions as done to derive (1) are standard features in the literature that followed.

Observe that case (a) occurs when the discount factor δ is sufficiently low. To see this, suppose payoffs are time invariant. Then the incentive constraint (1) can be written

$$\frac{\delta}{1-\delta}\left[q - c\left(q\right) - \pi - u\right] \geq c\left(q\right). \tag{4}$$

For a discount factor δ sufficiently close to 1, the left-hand side of (4) exceeds the right-hand side for $q = q^*$, so first-best performance can be implemented. But as the discount factor falls (the future is less important), eventually the left-hand side for $q = q^*$ must become less than the right-hand side and first-best performance can no longer be implemented. In this case the highest performance, $q^*\left(\delta\right)$, that gives payoffs on the Pareto frontier of self-enforcing stationary contracts has (4) hold with equality and solves

$$q^*\left(\delta\right) = \arg\max_q \delta q - c\left(q\right).$$

Observe that $q^*\left(\delta\right)$ decreases with the discount factor and that $\lim_{\delta \to 0} q^*\left(\delta\right) = 0$ which, in environments with low discount factors, effectively constrains the performance that can be supported by a relational contract. The discount factor may be low because of high interest rates, weak enforcement of contracts, or risk of government expropriation of assets. In economies where these are significant, relational contracts can be severely limited in what they can deliver.[4]

3.2 Renegotiation

Condition (1) in Theorem 1 requires that the parties jointly gain from continuation of the relationship over what they could obtain by reneging. This is a result about self-enforcing relational contracts, where self-enforcement is modeled as a sub-game perfect equilibrium. There is tension between the notion of "self-enforcement" and contract renegotiation that has a number of implications which connect this result to a rich informal literature on relational contracts (see for example MacNeil (1974), Goldberg (1980), Goetz and Scott (1981), and Bernstein (1992)).

Suppose that one party, either the purchaser or the supplier has breached the agreement to perform. Under the agreement parties are expected to move to an inefficient sub-game perfect equilibrium. A problem with this is that the breach event is like a sunk cost, and hence should be ignored by rational parties. In particular, this implies that, rather than punish each other, parties should agree to cooperate again. But then such behavior undoes the very action that is required to support cooperation in the first place — playing the inefficient one-shot Nash equilibrium if there is breach of contract terms. Thus an unfortunate consequence of outcomes that are *ex post* inefficient being renegotiated in the event that one party breaches the relational contract is that it makes such threats non-credible.[5] This has led to what is known as the problem of renegotiation proofness. MacLeod and Malcomson (1989) explicitly address this issue and show that if one applies the notion of renegotiation proofness due to Pearce (1992), the outcome is either an efficiency wage contract that gives the purchaser zero profits, or a bonus pay contract that leaves the supplier indifferent between the current relationship and the next best alternative (see Pearce and Stacchetti (2022) for a review of subsequent developments).

But there is a solution that we did not see at the time we wrote MacLeod and Malcomson (1989). In the same year we wrote that paper, van Damme (1989) showed that, when

[4]See McMillan and Woodruff (1999) for a discussion of relational contracts in a development context. MacLeod (2007) discusses the trade-off between relational contracts and formal contracts and their relevance for developing economies.

[5]The hypothesis that parties renegotiate inefficient allocations is one of the important ingredients of the property rights approach to the theory of the firm developed by Grossman and Hart (1986). See also Hart and Moore (1988)'s theory of incomplete contracts that builds upon the implications of efficient renegotiation, Schmidt and Schnitzer (1995) who show that the inclusion of enforceable contracts can undo informal contracts and the seminal work of Baker et al. (1994) integrating formal contracts into relational contract theory.

the Pareto frontier is linear, one can construct efficient renegotiation-proof equilibria for the repeated prisoners' dilemma game. The idea is simple. From Abreu (1988), one can show that any point on the Pareto frontier is supported by some sub-game perfect equilibrium. Hence, van Damme (1989) shows that for the prisoners' dilemma problem, if defection by party A occurs, parties play the equilibrium on the Pareto frontier that gives all the future gains to party B and vice versa. Levin (2003) shows that one can apply this argument to relational contracts and also extends the model to include moral hazard. The renegotiation-proofness argument is further strengthened in Goldlücke and Kranz (2013). (See also the work of Abreu and Sannikov (2013).) Thus, the main insight of the characterization of relational contracts holds despite efficient renegotiation when the Pareto frontier is linear (the parties are risk neutral). However, when the Pareto frontier is not linear (as when one of the parties is risk averse), there may not exist renegotiation-proof equilibria (Farrell and Maskin (1989)).

More generally, there are many principles that can be used to characterize the set of renegotiation-proof equilibria. Bergin and MacLeod (1993) conclude that this may help explain why the concept has not proven useful for explaining observed behavior.

3.3 Multiple Equilibria and Infinite Horizon

A common criticism of repeated game theory is the existence of multiple equilibria. In addition, in order to sustain cooperation it is assumed that there is an infinite horizon. Rubinstein (1991) addresses this issue, and observes that social phenomena are very complex. Hence the point of a model is to capture the essential features of the environment. In the case of Theorem 1, the main point is that, if parties believe they will meet again in the future and these meetings generate sufficient rent, they can ensure performance today. Crémer (1986) shows how participation of finitely-lived agents in organizations that are potentially infinitely-lived can sustain cooperation. But such a model does not deliver a unique solution, nor does it explain the empirical fact that parties cooperate in finitely repeated games.

Kreps et al. (1982) introduce a solution to both questions by supposing that there is a distribution of preferences, and some parties (irrationally) follow a strategy that leads to cooperation. In the case of the prisoners' dilemma, the assumption is that with some probability such parties plays tit-for-tat (see Axelrod (1981)). Bull (1987) shows that this idea can be used to establish the existence of self-enforcing relational contracts. While this appears to be a clean solution, it suffers from the problem that predictions depend upon the distribution of types in the population, which are not generally observable. Moreover, as Bergin and Lipman (1996) show, in dynamic models the equilibria can be very sensitive to small changes in the distribution of population characteristics. Miller and Watson (2022) discuss work in which they show that, by adding additional structure to the bargaining and renegotiation process, one can generate a unique equilibrium, a result that is reminiscent of Rubinstein (1982)'s model of bargaining.

3.4 Private Information

Building on Waldman (1984) and MacLeod and Malcomson (1989), MacLeod and Malcomson (1988) introduces some private information into a relational contract model that is motivated by evidence from the labor market.[6] Specifically, it studies the case in which the cost to the supplier of delivering performance (the supplier's *type*) is constant over time, not initially known to the purchaser and a purchaser observing shirking believes that the supplier is a higher cost type and either dismisses or demotes that supplier. This idea has some nice empirical implications.

In particular, MacLeod and Malcomson (1988) shows that, when no bonuses are paid, there exists an *equilibrium hierarchy* with a series of ranks (1 being the lowest), assumed observable to potential purchasers, in each of which there is an interval of types. Pay and performance increase with rank. Competition between purchasers ensures all supplier types who can be profitably employed will be. Moreover, suppose all supplier types start in rank 1 and are promoted to rank 2 if they perform sufficiently well, to rank 3 if they continue to perform sufficiently well, and so on. Then the purchaser can set performance criteria for promotion such that supplier types eventually sort themselves into equilibrium ranks. That is, all types reach the highest rank in which it is worthwhile for them to deliver the performance required to stay there.

In equilibrium, the performance criterion for promotion to a higher rank is higher than the performance criterion to remain in a rank. Moreover, implementing this promotion process is worthwhile for the purchaser. Those suppliers performing just well enough to remain in a rank generate no profit for the purchaser. But those striving for promotion perform better while being paid the same, so it is clearly profitable for the purchaser to set up the promotion system. A characteristic of the equilibrium hierarchy is that supplier

[6] Our recollection is that MacLeod and Malcomson (1988) was written after MacLeod and Malcomson (1989) but the publication process was quicker.

types are never fully revealed to the purchaser — all the purchaser eventually learns is in which of a set of intervals the supplier's type lies. For a more formal presentation, see Appendix B.

This model thus predicts an internal labor market with discrete ranks, and that individuals begin at the bottom and move up the ranks, as documented by Doeringer and Piore (1971). The model also predicts the Peter Principle (Peter and Hull (1969)) — individuals continue to be promoted until they reach their maximum rank, whereupon they continue to work, but at a lower productivity than individuals who move above them on the ladder.

In MacLeod and Malcomson (1989), the environment involves no uncertainty. Levin (2003) extends the model by adding shocks to the cost of effort by making the purchaser's profit in a period depend not directly on q_t but instead on a random variable y_t that is *i.i.d.* conditional on q_t in period t and distributed $F(y_t \mid q_t)$, with corresponding density function f. Moreover, before choosing q_t, the supplier observes a parameter of the cost function $\theta_t \in [\underline{\theta}, \overline{\theta}]$, referred to as the supplier's *type*, that is an *i.i.d.* draw for period t. The cost of delivering quality q_t, now denoted $c(q_t, \theta_t)$, is differentiable and increasing in θ_t, and satisfies the standard single crossing property, corresponding in the case it is twice differentiable to $\partial^2 c(q_t, \theta_t) / \partial q_t \partial \theta_t > 0$. The crucial difference from the models discussed above is that it is now optimal for performance to be a function $q(\theta)$ of type, with first-best performance $q^*(\theta)$ defined by

$$q^*(\theta) = \arg\max_q E_y[y \mid q] - c(q, \theta), \text{for all } \theta \in [\underline{\theta}, \overline{\theta}],$$

where E_y is the expectations operator over the random variable y. Levin (2003) considers two cases, both for a stationary environment. In one, the *hidden information* case, the purchaser observes q_t but not θ_t. In the other, the *moral hazard* case, the purchaser observes θ_t but not q_t.

For the hidden information case, Levin (2003) shows that, for a concave cost distribution, if the efficient performance $q^*(\theta)$ is unattainable for $\theta = \underline{\theta}$, an optimal $q(\theta)$ satisfies one of the following:

1. *pooling:* $q(\theta) = \tilde{q}$ for some $\tilde{q} < q^*(\theta)$, for all θ;

2. *partial pooling:* for some $\hat{\theta} \in (\underline{\theta}, \overline{\theta})$, $q(\theta) = \tilde{q} < q^*(\theta)$ for all $\theta \in [\underline{\theta}, \hat{\theta}]$ and $q(\theta) < q^*(\theta)$ and strictly decreasing for all $\theta \in (\hat{\theta}, \overline{\theta})$.

Thus either quality is the same for all types, or it is the same for an interval of the lowest cost types and decreasing in type for higher cost types. As in the basic model, quality is never above the efficient level $q^*(\theta)$ for *any* type. For the *moral hazard* case, Levin (2003) shows that, when the distribution F satisfies a convexity condition and the *monotone likelihood ratio property* that $\frac{\partial f(y|q)/\partial q}{f(y|q)}$ is increasing in y,

1. an optimal contract implements $q(\theta) \leq q^*(\theta)$ for all θ;

2. payments $P(\theta, y)$ are *one-step* in the sense that $P(\theta, y) = \underline{P}$ for all $y < \hat{y}(\theta)$ and $\overline{P} > \underline{P}$ for all $y \geq \hat{y}(\theta)$, where $\hat{y}(\theta)$ is the point at which the likelihood ratio switches from negative to positive as a function of y.

The intuition is that, with a risk-neutral supplier, it makes sense to use the limited incentive compatible bonuses in the way that gives the strongest incentives, so the maximum incentive is given for outcomes that are more likely to arise with high quality. Thus, as in the basic model, only one level of bonus is required. Levin (2003) also considers the case in which the purchaser observes y_t but the supplier does not, the case of a *subjective performance measure*, an issue that is developed in MacLeod (2003).[7]

3.5 Social Norms and Rent Allocation

As Schelling (1980), Binmore (1994) and Greif (1994) have observed, one can view game theory as providing a theory of social norms. While the idea of social norm is implicit in MacLeod and Malcomson (1989), it is far from obvious for the casual reader. There is a literature suggesting that wage formation and adjustment may be best understood in terms of such norms. Both Keynes and Marshall discuss this point when trying to explain rigid wages, while Akerlof (1980) explicitly links social norms to wage formation and unemployment.

The challenge is that starting from the *observation* of a social norm does not explain its source. This point is made beautifully in Skinner (1948), who shows that pigeons develop norms of behavior that have their source in luck and the rate at which food is delivered. MacLeod and Malcomson (1998) asks which of the multiple equilibria in MacLeod and Malcomson (1989) would be selected by a norm that prioritizes economic efficiency in a competitive matching market. That norm selects an equilibrium in which the rent from continuing the relationship that is required to sustain cooperation goes to the long side of the market. Intuitively, the number of matches is maximized by having all those on

[7]For further discussion of the results in Levin (2003), see Malcomson (2013).

the short side of the market matched, but then a defector on the short side can always obtain another match straightaway and so cannot obtain a rent from continuing the current match. The implication for a labor market is that the rent must go to employees in the form of efficiency wages (which implies that unemployment is involuntary) if there are more potential employees than jobs and to employers in the form of above-normal profits from filled jobs (which implies that vacancies are involuntarily unfilled and employees receive bonuses) if there are more jobs than potential employees.

Of course, which is the short side of a market is endogenous; MacLeod and Malcomson (1998) show how that depends on the cost of creating jobs. When these costs are low (for example in the case of service workers), then workers are on the short side of the market, and bonus pay is the equilibrium contract. This result is consistent with the observation that sales workers are paid bonus pay (MacLeod and Parent (1999)). Conversely, when the cost of job creation is high, as in capital intensive industries, then it is optimal to use an efficiency wage contract, a result that is consistent with the $5 dollar a day jobs offered by the Ford Motor Company (Raff and Summers (1987)).

This approach links the allocation of rents to observed contract form. It does not explain the evolution of rents over time. Ghosh and Ray (2023) reviews relational contract models with private information that provide a theory of trust building. Early in the relationship there is little trust, and hence performance is lower. Over time, the future value of the relationship rises, with the consequence that the efficiency of the relationship also rises over time. See Watson (1999) for a nice model with this feature that allows for contract renegotiation.

4 Our Subsequent Research

Great as it has been working together, Jim and Bentley's subsequent research on relational contracts has gone in different directions.

4.1 Jim's Subsequent Research

Jim's subsequent research on relational contracts has focused on two theoretical issues that have important practical implications: (1) how robust results are to shocks to the value of the supplier's performance to the purchaser and (2) whether there are equilibria other than that identified in MacLeod and Malcomson (1988) in which the supplier's persistent type is fully revealed.

Malcomson (2015b) extends the hidden information model in Levin (2003) to shocks to the value of the supplier's performance to the purchaser and shows that relational contracts favor pooling over purchaser types, thus making payment and performance independent of day-to-day changes in the purchaser's circumstances. This characteristic is encouraging for trying to understand supply relationships that, as documented by Carlton (1986), have prices that are relatively rigid over time despite apparent short-term changes in production conditions. But there remains research to be done to assess the extent to which those rigidities can genuinely be attributed to pooling across types that are private information. Malcomson (2021b) shows that such shocks induce risk-neutral parties to make investment choices as if they were risk-averse, a potential resolution to a macroeconomic puzzle in Bloom et al. (2018).

In MacLeod and Malcomson (1988), supplier types are never fully revealed to the purchaser. Malcomson (2016) asks whether there are other equilibria, including those that allow for bonuses, in which the supplier's type *is* fully revealed when types are continuous. The conclusion is that, if first-best performance is unattainable with a relational contract and continuation equilibria following full revelation of the supplier's type are on the Pareto frontier of payoffs for purchaser and supplier (which is one criterion for an equilibrium to be renegotiation proof), no perfect Bayesian equilibrium fully reveals supplier type. Although the formal argument is technical, the underlying reasoning is intuitive. If first-best performance is unattainable, the Pareto frontier is achieved with the highest performance that is incentive compatible for the supplier, see Section 3.1. This requires that the supplier is on the margin of indifference between delivering that performance and shirking. So, in the period before the supplier's type has been revealed, the supplier anticipates no strictly positive gain from continuing the relationship in the future. But the supplier can receive a strictly positive information rent by permanently delivering the lower performance intended for a higher-cost supplier type. An implication is that, for separation of a lower-cost type from a higher-cost type, the lower-cost type's performance must be discretely greater than that of the higher-cost type, which is not possible for every type in an interval. So full revelation of all types is not possible in any period that starts with an interval of pooled types no matter how many periods of separation there are. Malcomson (2016) shows that this remains the case even if separation in some period leaves no *interval* of pooled types. For a somewhat more formal presentation, see Appendix C.

Malcomson (2021a) characterizes the extent of separation of types for which first-best performance is not attainable if types are partitioned as finely as possible at each date. If the interval of types is sufficiently small for the given discount factor and performance cost function, no partitioning may be possible. For a somewhat larger interval of types, there is initial partitioning into two sub-intervals with no further partitioning thereafter. For an even larger interval of types, there is initial partitioning, with each resulting sub-interval further partitioned into two further sub-intervals in the following period but no further partitioning afterwards. Numerical illustrations for a constant elasticity cost of performance function show the maximum number of eventual partition intervals of types to be typically small (less than about a half dozen) despite supplier type being continuous. This is clearly not the final word on how much revelation will occur but deriving results in models of this type is far from straightforward. Since finding out information about the long-term characteristics of suppliers is fundamental to continuing productive relationships, it does, though, seem a worthwhile endeavor.

Although types are grouped in both MacLeod and Malcomson (1988) and Malcomson (2021a), the differences in what the groups provide are differences in performance level, not fundamental differences in the tasks to which different groups are best suited. An example of the latter would be the differences between the subcontractors to which Toyota supplies drawings of what is to be produced (*drawings supplied* subcontractors) and those who provide drawings to be approved by Toyota (*drawings approved* subcontractors) discussed by Asanuma (1989). Matching suppliers to tasks is important for productivity but also provides an additional dimension on which incentive compatibility can operate because there is a gain from doing it appropriately. Suppose for example, the purchaser commits to a fixed payment for drawings approved work that is higher than that for drawings supplied work by more than the difference in the cost of supply. Then a low-cost supplier has an additional incentive to reveal its type because, if promoted to drawings approved work, future profits will be greater. Moreover, the purchaser has an incentive to promote sufficiently low-cost types because it is more efficient to have them do drawings approved work. Gietzmann and Larsen (1998) have a formal model of this based on the promotion model in Fairburn and Malcomson (1994). The idea is developed in Fairburn and Malcomson (2001) into a model of the Peter Principle that is consistent with the evidence in Benson et al. (2019). Prendergast (1993) develops a model based on a similar underlying idea, though again not formally in a relational contract. Although these are finite horizon, and not relational contract, models, the underlying mechanism would seem to apply to the relational contract approach. But further work is needed to assess the implications for relational contracts of this important issue.

4.2 Bentley's Subsequent Research

MacLeod and Malcomson (1988, 1989) provide a framework within which private returns, social norms and rent allocation work together to allow productive exchange that would not occur otherwise. Bentley's subsequent research has focused upon extending these ideas to help explain observed economic institutions and performance. This work can be roughly divided into the following research questions. How do individuals learn to coordinate their behavior? Why is there conflict in many relationships, and how can it be mitigated? What is the role of compensation form in sustaining high performance?

4.2.1 How do individuals learn to cooperate and coordinate their behavior?

Standard, non-behavioral game theory is widely used to model relational contracts. However the voluminous experimental literature has shown that it has significant limitations as a model of observed behavior.[8] The extensive evidence shows that humans respond to incentives and that, when in a strategic situation, individuals take into account how others respond to their choices. However, these effects are imperfect. In any given experiment one frequently observes deviations from the predictions of game theory.

Relational contract theory allows one to view game theory through the lens of contract law. The legal definition of a contract explicitly entails a "meeting of the minds (see Kornhauser and MacLeod (2012)). Brandts and MacLeod (1995) study this question directly with a series of experiments in which the "meeting of the minds" is implemented by making recommendations to individuals on how they should play a game when there are multiple equilibria. If the Nash equilibrium concept indeed captures the idea of self-enforcing behavior, then individuals should follow play when the same Nash equilibrium is recommended to individuals before they play a game with monetary payoffs.

Brandts and MacLeod (1995) found that individuals follow the recommendation to play Nash equilibria when those equilibria are unique and stable. In particular, they follow the recommended play even when there exists another Nash equilibrium that Pareto dominates the recommendation. Hence, they show that in the absence of explicit discussion, the expectation that a Nash equilibria will be followed by a counter party, even an inefficient equi-

[8]For example, see Charness and Kuhn (2011) for a discussion of how experimental economics can be applied to labor economics.

librium, is sufficient to ensure parties follow the recommendations. However, in complex dynamic games, even when primed with sub-game perfect equilibrium recommendations, significant deviation from the equilibrium is observed.

This shows that by itself, the existence of efficient equilibrium play does not ensure that parties will follow the strategies entailed by this play. This result is consistent with the theoretical results in Bergin and MacLeod (1993) where it is shown that, without additional assumptions, there does not necessarily exist a well-defined notion of renegotiation-proof equilibrium.

Moreover, when the recommended play is unstable — more precisely the recommended strategy is weakly dominated by another, then deviation is observed. More generally, it is an empirical fact that in any experiment with human players there are often many deviations from Nash equilibrium play.[9] Evolutionary models provide a general approach to identifying stable equilibria (e.g. Taylor and Jonker (1978); Jovanovic (1982); Maynard Smith (1982)). In such models there is continual experimentation, and hence the prediction that individuals do not necessarily play the equilibrium outcome is consistent with learning and experimentation.

Carmichael and MacLeod (1997) introduce a market where individuals match and play a prisoners' dilemma game. As shown in Brandts and MacLeod (1995), communication can play an important role to help parties coordinate their actions. Carmichael and MacLeod study an environment in which parties can communicate at the beginning of a relationship before playing a prisoners' dilemma game, and then decide whether to continue the relationship or not. In such a case, it is optimal to promise to cooperate, and then cheat and leave the relationship. Thus the combination of communication and the freedom to exit a relationship does not necessarily lead to an efficient outcome.

The new insight of Carmichael and MacLeod (1997) is that an institution that requires some sort of *entry cost* can stabilize the environment on a cooperative equilibrium. An example of such an entry cost is gift exchange. Here a gift has a very specific economic meaning — the cost to the gift giver is higher than the benefit to the receiver (see Mauss (1990), Waldfogel (1993)). We show that, with gift exchange, cooperation is the unique evolutionarily stable outcome. The contribution to relational contract theory is to illus-

[9]The literature is too vast to cite here. However, Henrich et al. (2006) carry out a set of standardized experiments with different cultural groups. They find that there are systematic deviations from sub-game perfection that vary with the culture. This work shows that predicting individual behavior in a particular situation depends upon both the strategic features of the environment, and the culture of the individuals playing the game.

trate how the theory can be used to explain observed *costly* institutions that help initiate cooperation. See Sobel (2006) for an extension of these ideas to labor contracting.

4.2.2 Why is there conflict in many relationships, and how can it be mitigated?

A gift is a costly institution whose cost is paid at the beginning of a relationship. An alternative common institution is conflict that can occur either during or at the end of a relationship. Bentley has written a number of papers that explore the role of conflict in relational contracts that builds upon insights from the literature on asymmetric information (see Kennan and Wilson (1993); Ausubel et al. (2002)). MacLeod (2003) introduces a reduced form approach to thinking about relational contracts that uses a "relaxed budget constraint" to allow for conflict in a standard principal-agent model (Holmström (1979)). If one views bilateral exchange as a two-person team, then from Holmström (1982) and Eswaran and Kotwal (1984) we know an efficient contract does not exist.

The relaxed budget constraint is a short-hand way to add *ex post* conflict that can ensure cooperation.[10] Like gift exchange discussed above, it allows individuals to burn resources. The difference is that conflict entails the burning resources *after* trade, while gift exchange entails resource loss *before* trade. Potential conflict by the agent can ensure that the principal treats the agent *fairly* and rewards the agent whenever there is good performance. Potential conflict in this model corresponds to the separation decision and the loss of future earnings in MacLeod and Malcomson (1989).

Since MacLeod and Malcomson (1989) have symmetric information, there is never disagreement in equilibrium, and hence neither party suffers a loss along the equilibrium path. MacLeod (2003) and Levin (2003) provide formal models of subjective evaluation – namely the principal observes a private signal of employee performance. In this case, some form of conflict along the equilibrium path is a necessary condition for performance, and hence there can be disagreement regarding whether or not the agent has performed.[11] Hence, in equilibrium there can be conflict and a loss due to the existence of private information that is triggered by a disagreement regarding performance.[12]

[10] Frank (1988) is an early study that observes that potential *ex post* conflict can be used to enhance performance in relationships.

[11] Both models entail subjective evaluation. Levin (2003) begins with an explicit repeated game approach, and thus assumes parties are risk neutral to simplify the computation of the equilibrium. The benefit of using a relaxed budget constraint to model conflict is that MacLeod (2003) can allow agents to be risk averse. See the contribution by Thomas and Worrall (2022) on how risk aversion affects the design of relational contracts that smooth consumption over time.

[12] See Abreu et al. (1991) for a seminal analysis of asymmetric information in a repeated game and how

Observe that non-performance corresponds exactly to the legal notion of "contract breach", a central feature of contract law (Kornhauser and MacLeod (2012)). MacLeod (2007) shows that breach is not an exogenous event, but one that is defined by contract. For example, if a supplier produces a low quality good, under some contracts this might be called breach, which in turn may lead to conflict. However, as Levin (2003) shows, this is not efficient. For exactly the same economic relationship, one can respond to low quality with a payment from the seller to the buyer. In this case there is no breach since the possibility of the payment is anticipated in the contract. Breach would occur only if the seller did not make the payment as required under the contract.

This result shows that MacLeod and Malcomson (1989)'s invariance result that the efficiency of the relational contract is unaffected by the choice between bonus pay and an efficiency wage does not extend to the case with asymmetric information. MacLeod (2007), building upon the results of Abreu et al. (1991), shows that optimal breach design is a function of the characteristics of the commodity to be exchanged. For commodities for which low quality in equilibrium is rare, then a fixed price contract, with a penalty when there is failure, is efficient. However, for *innovative commodities* — commodities like research and development characterized by mostly failure (discoveries are infrequent), the optimal contract is a bonus payment whenever there is a success.

In work with Teck Yong Tan, we have extended this ideas to more general information structures (MacLeod and Tan (2022)). Most recently, MacLeod et al. (2020) reports results from an experiment where individuals are primed with an "organizational culture" that is intended to cue constrained (second-best) efficient behavior. In that "efficient" conflict leads to higher organizational performance. However, in contrast to the symmetric information games in Brandts and MacLeod (1995), it turns out to be quite difficult to elicit appropriate behavior from parties. This fits in with recent work in contract theory where misperception of performance can lead to conflict. This includes Hart and Moore (2007) on reference points, and Gibbons and Henderson (2012)'s work that emphasizes that importance of *clarity* for relational contracts.

4.2.3 What is the role of compensation form in sustaining high performance?

A key insight from the discussion in the previous subsection is that the presence of uncertainty in a relational contract implies that one can improve upon the performance of a

the type of information constrains the set of possible cooperative equilibria.

fixed wage relational contract with the addition of bonus pay. This section briefly discusses Bentley's empirical work on this question.

Bonus pay in MacLeod and Malcomson (1989) is a discretionary payment that is made after the principal has privately observed performance. Baker et al. (1994) emphasize that this type of pay plays a different role from what is normally considered to be *performance pay* — payments that are contractual obligations that vary with a verifiable signal of performance. Examples include payments as a function of amount of work produced, or as a function of total sales. MacLeod and Parent (1999) provides evidence on the relative importance of these types of payments in the US economy using the Panel Study for Income Dynamics (PSID), the National Longitudinal Survey of Youth (NLSY), the Quality of Employment Survey (QES) and the Current Population Survey (CPS). They find that in many occupations, particularly for highly skilled workers, more than 20% of the jobs entail some form of bonus pay. In particular, such jobs can be shown to entail more oversight by management. In contrast, with jobs in which there is a contractible performance pay, such as commission payments to sales persons, the use of such pay leads to less monitoring.

This result highlights the role of management for the subjective evaluation of workers, and the potential importance of relational contracts for understanding organizational performance, a point highlighted in Baker et al. (1994) (see also Baker et al. (2022)). In subsequent empirical work, Lemieux et al. (2012) show that the use of bonus pay can moderate employment shocks in a downturn. See also Kaur (2019). Yet there is a potentially interesting line of research that explores the effect of relational contracts to manage macroeconomic fluctuations (see MacLeod et al. (1994) for some early work along these lines). MacLeod and Parent (2014) use relational contract theory to divide jobs into contingent (performance is verifiable) and non-contingent (performance is non-verifiable) in order to bring together relational contract theory and the transactions cost approach of Williamson (2010). They find that adding transactions costs to relational contracts results in a richer model that can explain why individuals in more complex occupations have non-contingent pay and lower turnover.

A challenge one faces when trying to tease out the empirical implications of the theory with observational data is that one is never sure of identifying a causal effect (see Macchiavello and Morjaria (2022)). A solution is to use experimental evidence (Huffman (2022)). Falk et al. (2015), building on Brown et al. (2004), introduce an experiment to ask how contract institutions affect average performance. In particular, they find that, when there is employment protection, the introduction of bonus pay into a relational contract enhances performance. These results illustrate that with bonus pay one does not need to rely upon dismissal to provide performance incentives. In particular, the results are consistent with evidence in MacLeod and Parent (2014) that finds the use of such pay systems for highly skilled workers, which in turn leads to lower turnover and hence less loss of firm specific human capital. It is also consistent with the evidence Lemieux et al. (2009) that bonus pay occurs more frequently for highly skilled workers. Yet Lemieux et al. (2009) also show that bonus pay may be a significant contributor to inequality, which Breza et al. (2018) show may have a countervailing effect on performance.

5 Some Concluding Reflections

First, some reflections on what has been achieved by the formalization of relational contracts. Some earlier writers, such as Dore (1983), argued that relational contracts are inconsistent with economists' perceptions of allocative efficiency. As Dore (1983, p. 472) put it: "Any economist, at least any economist worth his neo-classical salt, would be likely to scoff at the idea (that relational contracts are conducive to economics efficiency). Just think, he would say, of the market imperfections, of the misallocation and loss of efficiency involved." That was, however, before the formal models discussed above were written. It should be apparent from the economics literature on relational contracts that research by economists on relational contracts is precisely about how the parties may achieve *more* efficient outcomes by making use of relational contracts in circumstances in which courts cannot be relied upon to enforce contractual agreements.

This formalization of relational contracts has established quite generally that obtaining more than minimal performance from an employee or supplier requires there to be a gain from continuing the relationship over what the parties can obtain by reneging. It has also tied down precisely how large this gain must be. One can think of that gain as an economic rent to continuing the relationship or as a cost to exit from it. In some cases, that gain arises naturally from investment in specific assets or skills, as discussed in Klein and Leffler (1981) and Malcomson (2015a). But where that is not the case, it has to be created specifically to enable cooperation. That is inconsistent with perfectly competitive markets in the future in which no party receives a payoff greater than the going market rate.

A major concern though is the extent to which the formal models capture what relational contracts in the real world are actually about (see Gil and Zanarone (2016) for a review of empirical work by economists). The findings of Macaulay (1963) provided an important motivation for economic research on relational contracts. However, models of the type presented here certainly do not capture all those findings. One role for relational contracts discussed by Macaulay (1963) is as a substitute for planning exchange relationships completely, a framework within which the relationship can be adjusted in the face of either unforeseen contingencies or inconsistent perceptions by the parties about what should happen in certain eventualities. That is not a feature of the models discussed above.

Game theory models make use of strategies that specify what actions the parties are to take for every history that can occur, and thus for every possible contingency. Moreover, equilibrium strategies are consistent in the sense that each party's strategy is a best response to the others'. Relational contracts in these models are a substitute for enforcement by courts, not a substitute for careful planning. There is certainly no modeling of unforeseen contingencies. That is an issue not only for theory — as Bernstein and Peterson (2021) point out, unforeseen contingencies seem to be a major reason for relational contracts breaking down in practice.

One approach to modeling unforeseen contingencies is as an investment problem, as in Bajari and Tadelis (2001) and Chakravarty and MacLeod (2009). There is a growing literature that combines relational contract theory with investment (Halonen (2002), Che and Sákovics (2004), Halac (2015) and Malcomson (2021b)) that might be extended to incorporate unforeseen contingencies. In an interesting paper, Englmaier and Fahn (2019) show that relational contract theory can explain over-investment into one's firm as a commitment device that increases the willingness of employees to work on behalf of the firm. Currie and MacLeod (2014) show that the standard model of decision making can be used to think about how the courts deal with unforeseen contingencies in contract law. It would be useful to see if this provides new insights to relational contract theory.

Another issue is the extent to which these models capture the notion of *trust* that is seen by many non-economists as central to relational contracts. Some of the economics literature interprets an equilibrium with performance above the level sustainable in a one-period relationship as involving trust because at least one party leaves itself vulnerable to exploitation by the other. As MacLeod (2007, p. 609) puts it: "In a relational contract, one party trusts the other when the value from future trade is greater than the one period gain from defection." In this spirit, Kvaløy and Olsen (2009) interpret an increase in the discount factor δ as an increase in trust because it increases the stake that the parties are prepared to risk. That interpretation of trust is not, however, universally accepted.

In the context of exchanges in which the gains from future dealings are sufficiently highly valued to induce cooperation, Sabel (1993, p. 1135) writes that: “it would be wrong to associate cooperation with trust at all, because cooperation results from continuous calculation of self-interest rather than a mutually recognized suspension, however circumscribed, of such calculation.” Sako (1992) develops different concepts of trust along these lines applicable to inter-firm relations in Britain and Japan. These writers view relational contracts as being about more than self-interested cooperation. Frydlinger and Hart (2021) perhaps provides a step in the right direction of capturing a formal role for trust. Recent empirical work by Bartling et al. (2021) shows that trust and contract enforcement can be complementary.

We conclude by emphasizing that relational contracts are a wonderful test bed for understanding economic performance that incorporates complex human behaviors. There is the element of mutual benefit when individuals expect to trade over an extended period of time, combined with the problem of managing such relationships when there is always the possibility of communication breakdown, conflict, and eventual termination of the relationship. A perennial goal for such relationships, and human society in general, is to have a happy ending!

References

Abreu, D. (1988). On the theory of infinitely repeated games with discounting. *Econometrica 56*(2), 383–396.

Abreu, D., P. Milgrom, and D. Pearce (1991, November). Information and timing in repeated partnerships. *Econometrica 59*(6), 1713–1734.

Abreu, D. and Y. Sannikov (2013). An algorithm for two-player repeated games with perfect monitoring. Technical report, Princeton University.

Akerlof, G. A. (1980, June). A theory of social custom of which unenployment may be one consequence. *Quarterly Journal of Economics 94*, 749–775.

Asanuma, B. (1989, March). Manufacturer-supplier relationships in Japan and the concept of relation-specific skill. *Journal of the Japanese and International Economies 3*(1), 1–30.

Ausubel, L. M., P. Cramton, and R. J. Deneckere (2002). Bargaining with incomplete information. In R. J. Aumann and S. Hart (Eds.), *Handbook of Game Theory*, Volume 3. Amsterdam, The Netherlands: Elsevier Science B. V.

Axelrod, R. (1981, June). The emergence of cooperation among egoists. *American Journal of Political Science 75*(2), 306–318.

Bajari, P. and S. Tadelis (2001). Incentives versus transaction costs: A theory of procurement contracts. *RAND Journal of Economics 32*(3), 387–407.

Baker, G., R. Gibbons, and K. J. Murphy (1994, November). Subjective performance measures in optimal incentive contracts. *The Quarterly Journal of Economics 109*(439), 1125–1156.

Baker, G., R. Gibbons, and K. J. Murphy (2022). From Incentives to Control to Adaptation: Exploring Interactions Between Formal and Relational Governance. *Forthcoming JITE*.

Bartling, B., E. Fehr, D. Huffman, and N. Netzer (2021, April). The complementarity between trust and contract enforcement. Technical Report 377, University of Zurich, Zurich, Switzerland.

Benson, A., D. Li, and K. Shue (2019). Promotions and the Peter principle. *Quarterly Journal of Economics 134*(4), 2085–2134.

Bergin, J. and B. L. Lipman (1996, July). Evolution with state-dependent mutations. *Econometrica 64*(4), 943–956.

Bergin, J. and W. B. MacLeod (1993). Efficiency and renegotiation in repeated games. *Journal of Economic Theory 61*(1), 42–73.

Bernstein, L. (1992). Opting out of the legal system: Extralegal contractual relations in the diamond industry. *The Journal of Legal Studies 21*(1), 115–157.

Bernstein, L. and B. Peterson (2021). Managerial contracting: A preliminary study. University of Chicago Law School.

Binmore, K. (1994). *Game Theory and the Social Contract. Volume 1. Playing Fair*. Series on Economic Learning and Social Evolution. Cambridge and London: MIT Press.

Blonski, M., P. Ockenfels, and G. Spagnolo (2011, August). Equilibrium selection in the repeated prisoner's dilemma: Axiomatic approach and experimental evidence. *American Economic Journal: Microeconomics 3*(3), 164–192.

Bloom, N., M. Floetotto, N. Jaimovich, I. Saporta-Eksten, and S. J. Terry (2018, May). Really uncertain business cycles. *Econometrica 86*(3), 1031–1065.

Bowles, S. (1985). The production process in a competitive economy: Walrasian, Neo-Hobbesian, and Marxian models. *American Economic Review 75*, 16–36.

Brandts, J. and W. B. MacLeod (1995). On the strategic stability of equilibria in experimental games. *Games and Economics Behavior 11*, 36–63.

Breza, E., S. Kaur, and Y. Shamdasani (2018). The morale effects of pay inequality. *Quarterly Journal of Economics 133*(2), 611–663.

Brown, M., A. Falk, and E. Fehr (2004, May). Relational contracts and the nature of market interactions. *Econometrica 72*(3), 747–780.

Bull, C. (1987, February). The existence of self-enforcing implicit contracts. *Quarterly Journal of Economics 102*(1), 147–159.

Carlton, D. W. (1986, September). The rigidity of prices. *American Economic Review 76*(4), 637–658.

Carmichael, H. L. (1985, December). Can unemployment be involuntary? comment. *American Economic Review 75*(5), 1213–1214.

Carmichael, H. L. and W. B. MacLeod (1997, July). Gift giving and the evolution of cooperation. *International Economic Review 38*, 485–509.

Chakravarty, S. and W. B. MacLeod (2009). Contracting in the shadow of the law. *Rand Journal of Economics 40*(3), 533–557.

Charness, G. and P. Kuhn (2011). Lab labor: What can labor economists learn from the lab? In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics, Volume 4*, Volume 4. Elsevier.

Che, Y.-K. and J. Sákovics (2004). A dynamic theory of holdup. *Econometrica 72*(4), 1063–1103.

Crémer, J. (1986, February). Cooperation in ongoing organizations. *Quarterly Journal of Economics 101*(1), 33–49.

Currie, J. M. and W. B. MacLeod (2014). Savage tables and tort law: An alternative to the precaution model. *The University of Chicago Law Review 81*(1), pp.53–82.

Doeringer, P. and M. Piore (1971). *Internal Labor Markets and Manpower Analysis*. Heath Lexington Books. Lexington, M.A., U.S.A.: Lexington Books.

Dore, R. (1983, December). Goodwill and the spirit of market capitalism. *British Journal of Sociology 34*(4), 459–482.

Englmaier, F. and M. Fahn (2019, November). Size matters: How over-investments relax liquidity constraints in relational contracts. *The Economic Journal 129*(624), 3092–3106.

Eswaran, M. and A. Kotwal (1984). The moral hazard of budget-breaking. *RAND Journal of Economics 15*(4), 578–581.

Fairburn, J. A. and J. M. Malcomson (1994, April). Rewarding performance by promotion to a different job. *European Economic Review 38*(3/4), 683–690.

Fairburn, J. A. and J. M. Malcomson (2001, January). Performance, promotion, and the Peter principle. *Review of Economic Studies 68*(1), 45–66.

Falk, A., D. Huffman, and W. B. MacLeod (2015). Institutions and contract enforcement. *Journal of Labor Economics 33*(3), 571–590.

Farrell, J. and E. Maskin (1989). Renegotiation in repeated games. *Games and Economic Behavior 1*, 327–360.

Frank, R. H. (1988). *Passions within Reason*. New York, NY, U.S.A.: W. W. Norton & and Company.

Freeman, R. D. and J. L. Medoff (1984). *What Do Unions Do?* New York, NY: Basic Books.

Friedman, J. W. (1971, Jan.). A non-cooperative equilibrium for supergames. *Review of Economic Studies 38*(113), 1–12.

Frydlinger, D. and O. Hart (2021, October). Overcoming contractual incompleteness: The role of guiding principles. Department of Economics Working Paper 640373, Harvard University.

Ghosh, P. and D. Ray (2023). The social equilibrium of relational arrangements. *Journal of Institutional and Theoretical Econonomics*.

Gibbons, R. and R. Henderson (2012, September-October). Relational contracts and organizational capabilities. *Organization Science 23*(5), 1350–1364.

Gietzmann, M. B. and J. G. Larsen (1998, September). Motivating subcontractors to perform development and design tasks. *Management Accounting Research 9*(3), 285–309.

Gil, R. and G. Zanarone (2016). New frontiers in empirical research on informal contracting. *Journal of Institutional and Theoretical Economics-Zeitschrift Fur Die Gesamte Staatswissenschaft 172*(2), 390–407.

Goetz, C. J. and R. E. Scott (1981). Principles of relational contracts. *Virginia Law Review 67*, 1089–1119.

Goldberg, V. P. (1980). Relational exchange: Economics and complex contracts. *American Behavioral Scientist 23*, 337–352.

Goldlücke, S. and S. Kranz (2013, July). Renegotiation-proof relational contracts. *Games and Economic Behavior 80*, 157–178.

Greif, A. (1994, October). 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.

Grossman, S. J. and O. D. Hart (1986, August). The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of Political Economy 94*(4), 691–719.

Halac, M. (SPR 2015). Investing in a relationship. *Rand J. Econ. 46*(1), 165–185.

Halonen, M. (2002, July). Reputation and the allocation of ownership. *Economic Journal 112*(481), 539–558.

Hart, O. D. and J. Moore (1988, July). Incomplete contracts and renegotiation. *Econometrica, 56*(4), 755–785.

Hart, O. D. and J. H. Moore (2007). Contracts as reference points. *Quarterly Journal of Economics 123*(1), 1–48.

Henrich, J., R. McElreath, A. Barr, J. Ensminger, C. Barrett, A. Bolyanatz, J. C. Cardenas, M. Gurven, E. Gwako, N. Henrich, C. Lesorogol, F. Marlowe, D. Tracer, and J. Ziker (2006, June). Costly punishment across human societies. *Science 312*(5781), 1767–1770.

Hirschman, A. (1970). *Exit, Voice, and Loyalty: Responses to Declines in Firms, Organizations, and States*. Cambridge, MA: Harvard University Press.

Holmström, B. (1979). Moral hazard and observability. *Bell Journal of Economics 10*(1), 74–91.

Holmström, B. (1982). Moral hazard in teams. *Bell Journal of Economics 13*, 324–40.

Huffman, D. (2022). Relational contracting: Complementarities with behavioral and experimental economics. *Forthcoming JITE*.

Jovanovic, B. (1982, May). Selection and evolution of industry. In *Econometrica*, pp. 649–670.

Kaur, S. (2019). Nominal wage rigidity in village labor markets. *American Economic Review 109*(10), 3585–3616.

Kennan, J. and R. Wilson (1993, March). Bargaining with private information. *Journal of Economic Literature 31*(1), 45–104.

Klein, B. and K. B. Leffler (1981, August). The role of market forces in assuring contractual performance. *Journal of Political Economy 89*(4), 615–641.

Kornhauser, L. and W. B. MacLeod (2012). Contracts between legal persons. In R. Gibbons and J. Roberts (Eds.), *Handbook of Organizational Economics*, Chapter 23, pp. 918–957. Princeton University Press.

Kreps, D. M., P. Milgrom, J. Roberts, and R. Wilson (1982). Rational cooperation in the finitely repeated prisoners-dilemma. *Journal of Economic Theory 27*(2), 245–252.

Kvaløy, O. and T. E. Olsen (2009, December). Endogenous verifiability and relational contracting. *American Economic Review 99*(5), 2193–2208.

Lemieux, T., W. B. MacLeod, and D. Parent (2009, February). Performance pay and wage inequality. *Quarterly Journal of Economics 124*(1), 1–49.

Lemieux, T., W. B. MacLeod, and D. Parent (2012, May). Contract form, wage flexibility and employment. *American Economic Review 101*(3), 526–531.

Levin, J. (2003, June). Relational incentive contracts. *American Economic Review 93*(3), 835–857.

Loasby, B. J. (1976). *Choice, Complexity and Ignorance*. Cambridge University Press.

Macaulay, S. (1963). Non-contractual relations in business: A preliminary study. *American Sociological Review 28*, 55–67.

Macchiavello, R. and A. Morjaria (2022). Relational Contracts: Recent Empirical Advancements and Open Questions. *Forthcoming JITE*.

MacLeod, W. (1985). A theory of conscious parallelism. *European Economic Review 27*(1), 25–44. Special Issue on Market Competition, Conflict and Collusion.

MacLeod, W. B. (1984, August). A theory of cooperative teams. Working paper, Université Catholique de Louvain, CORE Discussion Paper: 8441.

MacLeod, W. B. (1987). Behavior and the organization of the firm. *Journal of Comparative Economics 11*, 207–220.

MacLeod, W. B. (1988). Equity, efficiency and incentives in co-operative teams. *Advances in the Economic Analysis of Participatory and Labor Managed Firms 3*, 5–23.

MacLeod, W. B. (2003, March). Optimal contracting with subjective evaluation. *American Economic Review 93*(1), 216–240.

MacLeod, W. B. (2007, September). Reputations, relationships, and contract enforcement. *Journal of Economic Literature 45*, 595–628.

MacLeod, W. B. and J. M. Malcomson (1988, August). Reputation and hierarchy in dynamic models of employment. *Journal of Political Economy 96*(4), 832–854.

MacLeod, W. B. and J. M. Malcomson (1989, March). Implicit contracts, incentive compatibility, and involuntary unemployment. *Econometrica 57*(2), 447–480.

MacLeod, W. B. and J. M. Malcomson (1998, June). Motivation and markets. *American Economic Review 88*(3), 388–411.

MacLeod, W. B., J. M. Malcomson, and P. Gomme (1994). Labor turnover and the natural rate of unemployment: Efficiency wage versus frictional unemployment. *Journal of Labor Economics 12*(2), pp.276–315.

MacLeod, W. B. and D. Parent (1999). Job characteristics and the form of compensation. *Research in Labor Economics 18*, 177–242.

MacLeod, W. B. and D. Parent (2014). Transaction costs and the employment contract in the US economy. *Journal of Law, Economics and Organization 31*(Supp 1), 40–76.

MacLeod, W. B. and T. Y. Tan (2022, April). Chapter 10: Opportunism and conflict in agency relationships. In *Advanced Microenomics for Contract, Instituional and Organizational Economics*, pp. 299–233. Boston, MA: MIT Press.

MacLeod, W. B., V. Valle Lara, and C. Zehnder (2020, September). Worker empowerment and subjective evaluation: On building an effective conflict culture. Working Paper 27857.

MacNeil, I. R. (1974). The many futures of contracts. *Southern California Law Review 47*(688), 691–816.

Malcomson, J. M. (1981, December). Unemployment and the efficiency wage hypothesis. *Economic Journal 91*(364), 848–866.

Malcomson, J. M. (1984, June). Work incentives, hierarchy, and internal labor markets. *Journal of Political Economy 92*(3), 486–507.

Malcomson, J. M. (1986, October). Rank-order contracts for a principal with many agents. *Review of Economic Studies 53*(5), 807–817.

Malcomson, J. M. (1999). Individual employment contracts. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3B, Chapter 35, pp. 2291–2372. Amsterdam: Elsevier.

Malcomson, J. M. (2013). Relational incentive contracts. In R. Gibbons and J. Roberts (Eds.), *Handbook of Organizational Economics*, Chapter 25, pp. 1014–1065. Princeton, NJ: Princeton University Press.

Malcomson, J. M. (2015a, March). Relational contracts and specific training. *Research in Economics 69*(1), 51–62.

Malcomson, J. M. (2015b, July). Relational incentive contracts with productivity shocks. *Games and Economic Behavior 92*, 122–137.

Malcomson, J. M. (2016, January). Relational incentive contracts with persistent private information. *Econometrica 84*(1), 317–346.

Malcomson, J. M. (2021a, July). Grouping agents with persistent types. Department of Economics Discussion Paper 941, University of Oxford.

Malcomson, J. M. (2021b, July). Uncertainty, investment and productivity with relational contracts. Department of Economics Working Paper 940, University of Oxford.

Mauss, M. (1990). *The Gift*. New York: W. W. Norton.

Maynard Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge, UK.: Cambridge University Press.

McMillan, J. and C. Woodruff (1999). Interfirm relationships and informal credit in vietnam. *Quarterly Journal of Economics 114*(4), 1285–1320.

Miller, D. A. and J. Watson (2022). An active-contracting perspective on equilibrium selection in relational contracts. *Forthcoming JITE*.

Pearce, D. and E. Stacchetti (2022). Relational contracts: Reputation and renegotiation. *Forthcoming JITE*.

Pearce, D. G. (1992). Repeated games: Cooperation and rationality. In J.-J. Laffont (Ed.), *Advances in Economic Theory: Sixth World Congress*, Volume I, pp. 132–174. Cambridge, UK: Cambridge University Press.

Peter, L. J. and R. Hull (1969). *The Peter Principle*. William Morrow & Co.

Prendergast, C. (1993, May). The role of promotion in inducing specific human capital acquisition. *Quarterly Journal of Economics 108*(2), 523–534.

Raff, D. M. G. and L. H. Summers (1987, October). Did Henry Ford pay efficiency wages? *Journal of Labor Economics 5*(4, part 2), S57–S86.

Rubinstein, A. (1982, January). Perfect equilibrium in a bargaining model. *Econometrica 50*(1), 97–109.

Rubinstein, A. (1991, July). Comments on the interpretation of game theory. *Econometrica 59*(4), 909–924. 00129682 Econometric Society.

Sabel, C. F. (1993, September). Studied trust: Building new forms of cooperation in a volatile economy. *Human Relations 46*(9), 1133–1170.

Sako, M. (1992). *Prices, Quality and Trust: Inter-Firm Relations in Britain and Japan*. Cambridge, UK: Cambridge University Press.

Schelling, T. C. (1980). *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.

Schmidt, K. M. and M. Schnitzer (1995, May). The interaction of explicit and implicit contracts. *Economics Letters 48*(2), 193–199.

Shapiro, C. and J. E. Stiglitz (1984, June). Equilibrium unemployment as a worker discipline device. *American Economic Review 74*(3), 433–444.

Simon, H. A. (1979). Rational decision making in business organization. *American Economic Review 69*, 493–513.

Skinner, B. F. (1948). 'Superstition' in the pigeon. *Journal of Experimental Psychology 38*, 168–172.

Sobel, J. (2006, April). For better or forever: Formal versus informal enforcement. *Journal of Labor Economics 24*(2), 271–297.

Svenjar, J. (1982). On the theory of the participatory firm. *Journal of Economics Theory*.

Taylor, P. and L. Jonker (1978). Evolutionarily stable strategies and game dynamics. *Mathematical Biosciences 40*, 145–56.

Thomas, J. and T. Worrall (2022). Self-enforcing wage contracts redux. *Forthcoming JITE*.

van Damme, E. (1989, February). Renegotiation-proof equilibria in repeated prisoners' dilemma. *Journal of Economic Theory 47*(1), 206–17.

Waldfogel, J. (1993). The deadweight loss of Christmas. *American Economic Review 83*(5), 1328–1336.

Waldman, M. (1984). Job assignment, signaling, and efficiency. *RAND Journal of Economics 15*, 255–67.

Watson, J. (1999, March). Starting small and renegotiation. *Journal of Economic Theory 85*(1), 52–90.

Williamson, O. E. (2010, June). Transaction cost economics: The natural progression. *American Economics Review 100*, 673–690.

Williamson, O. E., M. L. Wachter, and J. E. Harris (1975, Spring). Understanding the employment relation: the analysis of idiosyncratic exchange. *Bell Journal of Economics 6*(1), 250–278.

Appendices (for online)

Appendix A Sketch proof of Theorem 1

In a potentially on-going relationship, the supplier's future payoff from t on for a relationship started at τ if both parties stick to their agreement consists of the payment $p_t + b_t$ less the cost of supply $c_t(q_t)$ in period t plus the expected payoff from the future. It can be written

$$U_t = p_t + b_t - c_t(q_t) + \delta_t U_{t+1}, \quad \text{for all } t \geq \tau. \tag{A.1}$$

If, on the other hand, the supplier delivers performance in breach of the agreement, the worst that can happen is that the relationship comes to an end with payment of p_t, but without payment of the bonus b_t, because the supplier can always choose to quit at the end of period t. Let $\tilde{U}_{t+1}$ denote the supplier's payoff from $t+1$ on in those circumstances. Thus, by setting $q_t = 0$, the supplier can obtain a payoff from t on of no less than $p_t + \delta_t \tilde{U}_{t+1}$, for all $t \geq \tau$. From this and (A.1), it is certainly better for the supplier to set $q_t = 0$ than to stick to any agreement with $q_t > 0$ unless the following incentive compatibility condition holds

$$\delta_t \{\text{future gain to supplier}\} \equiv \delta_t \left(U_{t+1} - \tilde{U}_{t+1}\right) \geq c_t(q_t) - b_t, \quad \text{for all } t \geq \tau. \tag{A.2}$$

The interpretation of this condition is that the future gains from having the relationship continue at the end of period t must exceed the cost $c_t(q_t)$ of complying less the current period incentive to supply provided by the bonus b_t. Moreover, for the supplier to continue in the relationship when there has been no breach, the following individual rationality condition must also hold

$$U_t \geq \bar{U}_t, \quad \text{for all } t \geq \tau, \tag{A.3}$$

where $\bar{U}_t$ denotes the supplier's payoff from t on if the relationship ends for reasons unrelated to the performance of either party. In principle, $\bar{U}_t$ need not be the same as $\tilde{U}_t$.

The purchaser's future payoff from t on for a relationship started at τ if both parties stick to their agreement consists of the value q_t derived from supply less the payment $p_t + b_t$ in period t plus the expected payoff from the future. It can be written

$$\Pi_t = q_t - p_t - b_t + \delta_t \Pi_{t+1}, \quad \text{for all } t \geq \tau. \tag{A.4}$$

If the purchaser does not pay the agreed bonus, the worst that can happen is that the relationship comes to an end because the purchaser can always choose to quit at the end of period t. Let $\tilde{\Pi}_{t+1}$ denote the purchaser's payoff from $t+1$ on in those circumstances. Thus, by setting $b_t = 0$, the purchaser can obtain a payoff from t on of no less than $q_t - p_t + \delta_t \tilde{\Pi}_{t+1}$. Thus it is certainly better for the purchaser to pay no bonus than to stick to any agreement with $b_t > 0$ unless the following incentive compatibility condition holds

$$\delta_t \{\text{future gain to purchaser}\} \equiv \delta_t \left(\Pi_{t+1} - \tilde{\Pi}_{t+1}\right) \geq b_t, \quad \text{for all } t \geq \tau. \tag{A.5}$$

Moreover, for the purchaser to continue in the relationship when there has been no breach, the following individual rationality condition must clearly hold

$$\Pi_t \geq \bar{\Pi}_t, \quad \text{for all } t \geq \tau, \tag{A.6}$$

where $\bar{\Pi}_t$ denotes the purchaser's payoff from t on if the relationship ends for reasons unrelated to the performance of either party. In principle, $\bar{\Pi}_t$ need not be the same as $\tilde{\Pi}_t$.

Conditions (A.2) and (A.5) can be added to give the pooled incentive compatibility condition (1). It is clear from the way they have been derived that (1), (A.3) and (A.6) (and hence (2)) are necessary conditions for the parties to stick to any agreement with $q_t > 0$, that is, for the agreement to be self-enforcing. Following the argument in MacLeod and Malcomson (1989), they are also sufficient. Formally, provided an agreement satisfies (1), (A.3) and (A.6), there exist strategies implementing the agreement that form a sub-game perfect equilibrium. Hence, (1), (A.3) and (A.6) provide a complete characterization of sub-game perfect equilibria with qualities strictly greater than 0.

Appendix B Derivation of equilibrium hierarchy

In MacLeod and Malcomson (1988), the environment is stationary with the cost of performance $c(q)/\theta$, where $\theta \in \left[\underline{\theta}, \overline{\theta}\right]$ with $\underline{\theta} > 0$ is the supplier's type that is known to the supplier but not the purchaser, and no bonuses, so the only payment is the fixed payment. Because the supplier has private information, the equilibria considered are perfect Bayesian equilibria. Suppose the purchaser were to group supplier types into a set of ranks r (1 being the lowest), assumed observable to potential purchasers, with payment $p^r > p^{r-1}$, minimum performance $q^r > q^{r-1}$ to stay in r and those in rank r who deliver performance below q^r demoted to $r-1$ (or dropped as a supplier and used by another purchaser in the equivalent of rank $r-1$). The expected utility from staying in rank r for ever is then

$$U^r(\theta) = p^r - \frac{c(q^r)}{\theta} + \delta U^r(\theta), \quad \text{for all } r \geq 1, \text{ all } \theta, \tag{B.1}$$

which corresponds to (A.1) in Appendix A with the bonus set to zero. Because other potential purchasers observe a supplier's rank, the payment p^r for rank r is determined by the zero-profit condition $p^r = q^r$, so (B.1) can be solved to give

$$U^r(\theta) = \left[q^r - \frac{c(q^r)}{\theta}\right] / (1-\delta), \quad \text{for all } r \geq 1, \text{ all } \theta. \tag{B.2}$$

It is convenient to define $U^0(\theta)$ as the future payoff to type θ if dropped from the lowest rank (rank 1) permanently with payoff $U^0(\theta) = u/(1-\delta)$.

Consider which types θ will perform well enough to stay in rank r if they reach it. The incentive compatibility condition for this, corresponding to (A.2) with no bonus, is

$$\delta\left[U^r(\theta) - U^{r-1}(\theta)\right] \geq \frac{c(q^r)}{\theta}, \quad \text{for all } r \geq 1, \text{ all } \theta. \tag{B.3}$$

Define θ^r as the lowest value of θ that satisfies this condition. Substitution from (B.2), use of the convention that $q^0 = 0$ and $c\left(q^0\right) = -u$, and re-arrangement gives

$$\theta^r = \frac{1}{q^r - q^{r-1}}\left[\frac{c(q^r)}{\delta} - c\left(q^{r-1}\right)\right], \quad \text{for all } r \geq 1. \tag{B.4}$$

An *equilibrium hierarchy* is a triple (q^r, p^r, θ^r) that satisfies this condition and the zero-profit requirement that $p^r = q^r$ for each r. Competition between purchasers ensures all supplier types who can be profitably employed will be. In particular, θ^1 must be at the lowest value that is incentive compatible, so q^1 is determined to minimize the right-hand side of (B.4) for $r = 1$, given the conventions $q^0 = 0$ and $c\left(q^0\right) = -u$. This determines q^1 and θ^1. Given q^1 and θ^1, q^2 is determined to minimize the right-hand side of (B.4) for $r = 2$. This determines q^2 and θ^2 and so on up to rank R such that $\theta^{R+1} > \overline{\theta}$, the lowest-cost type. This iterative process determines a unique equilibrium hierarchy.

Suppose all supplier types start in rank 1 and are promoted to rank 2 if they perform sufficiently well, then to rank 3 if they continue to perform sufficiently well, and so on. Suppose the purchaser sets promotion criterion $\hat{q}^r$ for promotion from rank r to $r+1$. To sort properly, it must be that types $\theta \geq \theta^{r+1}$ perform sufficiently well for promotion. They will do that only if meeting the promotion criterion $\hat{q}^r$ to gain promotion to rank $r+1$ and then staying there has a higher payoff than staying in rank r. That is, only if

$$p^r - \frac{c(\hat{q}^r)}{\theta} + \delta U^{r+1}(\theta) \geq U^r(\theta), \quad \text{for } \theta \geq \theta^{r+1}, r \geq 1.$$

For selection to occur, it must also be that types $\theta < \theta^{r+1}$ do not find it worthwhile to deliver sufficiently high performance for promotion so

$$p^r - \frac{c(\hat{q}^r)}{\theta} + \delta U^{r+1}(\theta) < U^r(\theta), \quad \text{for } \theta < \theta^{r+1}, r \geq 1.$$

With θ continuous, these two conditions imply the sorting condition

$$p^r - \frac{c(\hat{q}^r)}{\theta^{r+1}} + \delta U^{r+1}\left(\theta^{r+1}\right) = U^r\left(\theta^{r+1}\right), \quad \text{for all } r \geq 1. \tag{B.5}$$

Substitution for $U^r\left(\theta^{r+1}\right)$ from (B.1) gives

$$\delta\left[U^{r+1}\left(\theta^{r+1}\right) - U^r\left(\theta^{r+1}\right)\right] = \frac{c(\hat{q}^r)}{\theta^{r+1}} - \frac{c(q^r)}{\theta^{r+1}}, \quad \text{for all } r \geq 1. \tag{B.6}$$

Because θ^{r+1} is, by definition, the lowest θ for which the incentive compatibility condition for the supplier (B.3) is satisfied for rank $r+1$, it follows from (B.3) that

$$\delta\left[U^{r+1}\left(\theta^{r+1}\right) - U^r\left(\theta^{r+1}\right)\right] = \frac{c\left(q^{r+1}\right)}{\theta^{r+1}}, \quad \text{for all } r \geq 1. \tag{B.7}$$

Equating the right-hand sides of (B.6) and (B.7) gives

$$c(\hat{q}^r) = c\left(q^{r+1}\right) + c(q^r), \quad \text{for all } r \geq 1,$$

which implies $\hat{q}^r > q^{r+1}$, for $r \geq 1$. So the performance required to be promoted from rank r to rank $r+1$ is higher than that required to stay in rank $r+1$ once there. This process eventually results in a supplier with θ between θ^r and θ^{r+1} being sorted into rank r. That is, all types reach the highest rank in which it is worthwhile for them to perform well enough to stay there. Moreover, implementing this promotion process is worthwhile for the purchaser. Those performing just well enough to stay in a rank generate no profit for the purchaser. But those striving for promotion perform better while being paid the same, so it is clearly profitable for the purchaser to set up the promotion system.

Appendix C Full revelation of supplier type

Malcomson (2016) uses the more general cost of performance $c(q,\theta)$ for c decreasing in the supplier's type θ. To illustrate the result there, consider the case with performance cost $c(q)/\theta$ and the supplier's future payoff if separating from the purchaser is zero. Consider type θ in an interval of types $[\theta',\theta'']$ with $\theta''>\theta'$ pooled at the beginning of period t to be fully revealed at t by performing at $q_\tau(\theta)$ and receiving payments $p_\tau(\theta)$ for $\tau\geq t$, whereas θ' is to perform at $q_\tau(\theta')$ and receive payments $p_\tau(\theta')$. With type θ fully revealed at t and so indifferent between delivering Pareto frontier performance $q_{t+1}(\theta)$ at $t+1$ and shirking, θ would have future payoff at the beginning of period t from delivering $q_\tau(\theta)$ of $p_t(\theta)-c(q_t(\theta))/\theta$ and from instead choosing $q_t(\theta')$ of $\sum_{\tau=t}^{\infty}\delta^{\tau-t}(p_t(\theta')-c(q_t(\theta'))/\theta)$. Type θ' would have future payoff from delivering $q_t(\theta')$ of $\sum_{\tau=t}^{\infty}\delta^{\tau-t}(p_t(\theta')-c(q_t(\theta'))/\theta')$ but, by delivering $q_t(\theta)$ at t and then quitting the relationship at $t+1$ of $p_t(\theta)-c(q_t(\theta))/\theta'$. The standard mechanism design condition for separation at t is that the difference in payoff between θ and θ' from choosing $q_t(\theta)$ must be at least as great as the difference in payoff from choosing $q_\tau(\theta')$ which, in this case corresponds to

$$-c(q_t(\theta))\left(\frac{1}{\theta}-\frac{1}{\theta'}\right)\geq-\left[c\left(q_t\left(\theta'\right)\right)+\sum_{\tau=t+1}^{\infty}\delta^{\tau-t}c\left(q_t\left(\theta'\right)\right)\right]\left(\frac{1}{\theta}-\frac{1}{\theta'}\right)$$

or, with $\theta>\theta'$,

$$c(q_t(\theta))\geq c\left(q_t\left(\theta'\right)\right)+\sum_{\tau=t+1}^{\infty}\delta^{\tau-t}c\left(q_t\left(\theta'\right)\right).$$

The summation term in this is the information rent θ can obtain by choosing the performance intended for θ', which is necessarily bounded above zero. So the revelation condition requires $q_t(\theta)$ greater than $q_t(\theta')$ by a discrete amount, which is not possible for every θ in the interval $[\theta',\theta'']$ with q_t bounded. Full revelation of all types $\theta\in[\theta',\theta'']$ at t is, therefore, not possible. It follows by the same argument that it is not possible in any period that starts with an interval of pooled types no matter how many periods of separation there are. Malcomson (2016) shows that this remains the case even if separation at t leaves no *interval* of pooled types.