Econometrica, Vol. 57, No. 2 (March, 1989), 447–480

IMPLICIT CONTRACTS, INCENTIVE COMPATIBILITY, AND INVOLUNTARY UNEMPLOYMENT

BY W. BENTLEY MACLEOD AND JAMES M. MALCOMSON[1]

This paper considers the enforceability of employment contracts when employees' performance cannot be verified in court so that piece-rate contracts are not legally enforceable. Part I shows that there exists a variety of self-enforcing implicit contracts, modelled as perfect equilibria in a repeated game, and characterizes all the wage and performance outcomes that can be implemented. Implementation requires a strictly positive surplus from employment, the form of the contract depending on how this surplus is divided between firm and employee. Piece-rate contracts, and contracts with an informally agreed bonus, can be made self-enforcing but the use of severance pay and bonding does not extend the set of implementable allocations. The resulting contracts resemble actual labor contracts more than do the contracts in standard principal-agent models.

Part II analyses market equilibrium with these contracts, also modelled as perfect equilibria in a repeated game, and shows that many such equilibria exist. Unfilled vacancies and unemployed workers can co-exist despite the existence of contracts that are potentially mutually beneficial. For those jobs that are filled, any division of the potential surplus is possible so that the market can have, at the same time, involuntary unemployment and vacancies that are unfilled despite filled jobs earning positive profits. As a criterion for selecting equilibria, a notion of renegotiation proofness is applied. Then either all workers are employed or all jobs filled but any division of the potential surplus is still possible. The paper explores what further restrictions on beliefs give rise to a Walrasian outcome, in which all the potential surplus goes to the short side of the market, and to an efficiency wage type outcome, in which the potential surplus goes to the long side.

KEYWORDS: Implicit contracts, involuntary unemployment, moral hazard.

PART I: SELF-ENFORCING CONTRACTS

1. INTRODUCTION

MORAL HAZARD PERVADES EMPLOYMENT. It is hard to think of jobs in which there is no scope for moral hazard. Yet principal-agent models, the standard tool of economists for analyzing such situations, generate contracts quite unlike most actual labor contracts. As Hart and Holmstrom (1987) note, "the extreme sensitivity to informational variables that comes across from this type of modelling is at odds with reality."

Standard principal-agent models have payment based on some measure of performance. In effect, these are piece-rate contracts. To be legally enforceable, the courts must be able to verify that performance. But finding verifiable measures of performance poses practical difficulties. There is now an increasing literature, stemming from Alchian and Demsetz (1972) through Williamson, Wachter, and Harris (1975) to Bull (1987), Malcomson (1981, 1984, 1986) and

[1] We wish to thank Richard Arnott, John Moore, two anonymous referees and participants in numerous seminars for helpful comments on earlier versions of the paper and CORE and the Départment des Sciences Economiques at the Université Catholique de Louvain for research facilities. We also gratefully acknowledge financial support from, respectively, the SSHRC in Canada and the Economic and Social Research Council (Award Number F00242005) in the U.K.

Shapiro and Stiglitz (1984), recognizing that this poses serious limitations on contracts. Most of that literature assumes that, without verifiable performance, contracts must be of the termination type, in which an employee receives an explicitly agreed wage as long as performance in the past has been satisfactory but is dismissed if it has not. Such contracts look more like actual labor contracts than those in standard principal-agent models. They also have important implications for labor markets. Shapiro and Stiglitz (1984), for example, conclude that they necessarily result in involuntary unemployment. Yet, piece-rate contracts are actually used, in many cases even where the arrangements for recording output are sufficiently informal that verifying it in court would be difficult. Moreover, in some jobs a substantial part of pay takes the form of a performance-related bonus (in effect, an informal piece-rate) which employees expect to be paid, and firms actually pay, despite there being no legal obligation to do so.

In the present paper, instead of imposing a particular contract form, we ask what contracts, and hence market outcomes, can be sustained when the only legally verifiable pieces of information are money payments and whether or not a person is employed by a firm. Implicit contracts have an important role in this. By an implicit contract, we mean one in which some elements are not legally enforceable, although typically there are also some explicit elements that are. Since they are not legally enforceable, implicit contracts will be carried out only if both parties wish it (that is, if they are self-enforcing) but that does not make them valueless. Indeed, termination contracts have their implicit element, namely that the employee will perform satisfactorily if employed and that the employer will continue the contract if performance is satisfactory, or terminate it if not. For the contract to be effective, both need to know that the other is behaving in this way. It is the need to make this implicit contract self-enforcing that gives rise to the involuntary unemployment equilibrium of Shapiro and Stiglitz (1984). But, as we show, this is not the only type of implicit contract that can be made self-enforcing.

In Part I of this paper, we provide a complete characterization of the sequences of wage and performance levels that can be implemented by implicit contracts between a firm and an employee and show what combinations of implicit and explicit contracts implement these. To do this we model employment as a repeated game. Following Aumann's (1974) insight of viewing equilibrium points of games as self-enforcing agreements, it is natural to model self-enforcing implicit contracts as perfect equilibria of this game. Since our characterization is complete, it is not open to the critique Carmichael (1985) makes of Shapiro and Stiglitz (1984) that the results arise from the assumption of a particular form of contract. The fundamental requirement for an implicit contract to be self-enforcing is that there exist sufficient economic surplus from continuing it over what the parties can jointly get if it is terminated. Given sufficient total surplus, however, any division of that surplus between the firm and the employee can be supported by some self-enforcing contract, the form of the contract depending on how the surplus is divided. Piece-rate contracts, and contracts with informally agreed bonuses, can be made self-enforcing. Explicit contracts with severance pay, however, can achieve no more than those without it nor, despite their popularity with theorists as a recipe for moral hazard, can contracts with bonding. Indeed, with suitable timing of payments, the absence of any legally enforceable agreement whatever imposes no further limit on what can be achieved. These results are an important step in understanding labor contracts. Piece rates and bonuses occur in practice despite performance not being verifiable. Severance pay and bonding that are negotiated as part of a contract, as opposed to being legally imposed or voluntarily offered by one party, are not that common. And written employment contracts are, in terms of the history of markets, a relatively recent phenomenon.

In Part II of the paper, we consider market equilibrium in a one-sector general equilibrium version of the model with a given number of workers and a given number of jobs. The market is formalized as an explicit game and we characterize its perfect equilibria. In the spirit of the "Folk Theorem" results for repeated games, we show that any self-enforcing combination of firms' profits and workers' utilities that gives all agents at least as much as if they did not participate in the market can be an equilibrium. In particular, there is no guarantee that all potential matches between workers and jobs will be made. Unemployed workers and unfilled vacancies can co-exist despite the existence of mutually beneficial wage and performance sequences that can be implemented by self-enforcing contracts. In effect, if agents do not believe that other agents will carry out their part of an implicit contract, they will not form matches even when those matches could be mutually beneficial. Moreover, employed workers can receive strictly greater lifetime utility than unemployed workers, in which case unemployment is involuntary, and filled vacancies can earn strictly greater profits than unfilled ones so that, using an obvious terminology, vacancies are involuntarily unfilled.

In these results, as in the standard "Folk Theorem" results, agents who deviate from an equilibrium are punished by other agents refusing to cooperate with them in the future. But, because the potential gains from cooperation remain, there is an incentive to renegotiate contracts to reap those gains even if some agent has deviated in the past. For this reason, it has been doubted whether such punishments are really credible. But, if agents are prepared to renegotiate, they should recognize that those punishments are not credible and take account of that in their original contracts. To be self-enforcing, a contract would then need to be proof against such renegotiation. We therefore consider which market equilibria are renegotiation proof, in a sense related to Pearce (1987). For this, either all workers must be employed or all jobs filled but there is still a potential surplus that does not necessarily go to agents on the short side of the market. Who gets that surplus depends on the beliefs (in the sense of Aumann (1987)) that agents have about the choice of strategies by other players. We explore two sets of beliefs, one giving rise to an outcome that is Walrasian in the sense that agents on the short side of the market get all the surplus, the other giving rise to an efficiency wage type outcome in which they get none of the surplus. In the latter case, when there are more workers than jobs all vacancies are filled but the unemployed may get lower lifetime utility than the employed so that unemployment is involuntary. This provides a rigorous theoretical basis for the type of efficiency wage unemployment equilibrium of Shapiro and Stiglitz (1984). But conversely, when there are more jobs than workers, all workers are employed, the surplus goes to firms, and there are involuntarily unfilled vacancies. So involuntary unemployment is not a necessary consequence of these beliefs. Perhaps the most important message to draw from this is that beliefs are crucial in determining equilibria in markets of this sort. To understand those markets better, economists need to understand the beliefs of the agents operating in them.

The earlier literature, for example Azariadis (1975), based the use of implicit contracts on the complexity of writing state-contingent contracts to cover every eventuality. Holmstrom (1981) and Thomas and Worrall (1988) have considered what implicit contracts are self-enforcing in these models. Simon (1957) emphasized the importance of bounded rationality, the inability of agents to foresee all the relevant contingencies, in limiting the extent to which contracts can be made explicit. Our approach is closer to that of Grossman and Hart (1981), in which it is the inability to verify information that limits explicit contracts. Their model, however, is based on risk sharing under adverse selection, ours on moral hazard. Here we abstract from risk sharing in order to focus on the essential consequences of nonverifiability. Bull (1987) also investigates self-enforcing contracts in a model in which performance is not verifiable but is concerned with the role of firms' reputations in ensuring the existence of efficient self-enforcing contracts when workers have finite lifetimes and when a firm and a worker, once together, cannot separate to form other matches. We have taken up the issues of reputations, in this case on the employee's side, in MacLeod and Malcomson (1988). In MacLeod and Malcomson (1989) we explore the implications of the present framework for the relationship between layoffs, labor turnover, and unemployment when the productivity of jobs fluctuates over time.

The plan of the present paper is as follows. Part I consider contracts between a firm and an employee. Section 2 sets out the basic assumptions of the model. Section 3 specifies the precise nature of the game they play and its relation to implicit contracts. In Section 4, we characterize the sequences of wages and performance levels that can be implemented by self-enforcing contracts for given default profit and utility levels that the firm and the employee would receive if they did not stay together. We also establish which contracts implement those sequences. Part II turns to market equilibrium, which determines the default profit and utility levels. In Section 5, we consider equilibria that are subgame perfect, in Section 6 equilibria that are also renegotiation proof.

2. THE MODEL

Consider a single firm and one of the many workers employed by it. Both are infinitely lived.[2] If employed by the firm in period $t \geqslant 1$, the value of the worker's

[2] The assumption of infinite lives avoids the problem, common in finitely repeated games, of agents inevitably defaulting on a nonbinding agreement in the final period so that any such agreement unravels to a Nash equilibrium of the corresponding one-period game. When workers have finite lives, this problem can be overcome as long as sufficient information about the actions of firms is conveyed from one generation of workers to the next; see Bull (1987). We conjecture that this is also the case for certain other types of reputation equilibria but to set up the model in this way adds greatly to the complexity of the analysis and has not therefore been done here.

output, net of all nonlabor costs, is $y(p_t)$, where $p_t \geqslant 0$ is a measure of the worker's performance in period t. This output is not verifiable in court.

ASSUMPTION 1: *$y(p_t)$ is continuous, strictly increasing, and concave for $p_t \geqslant 0$, with $y(0) = -k < 0$.*

In this, k is a fixed cost of employment incurred by the firm only if it employs a worker. Let w_t be the wage paid to the worker in period t and define the allocation x resulting from a contract by

$$(2.1) \qquad x \equiv \{(w_1, p_1), (w_2, p_2), \ldots, (w_T, p_T), w_{T+1}\} \in X,$$

where T is the last date at which the worker is employed by the firm and X is the set of all possible allocations, assumed to be bounded sequences. The term w_{T+1} is included to allow for the (present discounted value of) any severance pay or pension paid after the worker stops being employed by the firm. Let δ denote the discount factor and $\Pi_t^0 \geqslant 0$ the firm's expected discounted future profits from that production function from period t on if it employs the worker only to period $t-1$ or, in the case of $t = 1$, never employs the worker. Note that, since the firm incurs the fixed cost only if it employs the worker, $\Pi_t^0 \geqslant \delta \Pi_{t+1}^0$. Then, for the allocation x, the present discounted value at time t of the firm's profits from period t on is given by (with the summation taken as zero for $t > T$)

$$(2.2) \qquad \Pi_t(x) = \sum_{\theta=t}^{T} [y(p_\theta) - w_\theta]\, \delta^{\theta-t} + \delta^{T-t+1}(-w_{T+1} + \Pi^0{}_{T+1}) \qquad (t = 1, \ldots, T+1).$$

The worker's performance depends on effort. Like output, performance is not verifiable. In practice, a firm's measure of a worker's performance may be as concrete as a count of the physical units produced or as nebulous, and as subject to error, as an informal assessment by a supervisor. To keep the model simple, we assume here that there is no measurement error, so measured performance is the same as actual performance, which is thus known to both the firm and the worker. The worker's utility can then be written as a function of income and performance. Let $U_t^0 \geqslant 0$ denote the expected lifetime utility of the worker from period t on if employed by the firm only to period $t-1$ or, if $t = 1$, never employed by the firm.

ASSUMPTION 2: *The worker's lifetime utility from period t on is given by*

$$(2.3) \qquad U_t(x) = \sum_{\theta=t}^{T} [w_\theta - v(p_\theta)]\, \delta^{\theta-t} + \delta^{T-t+1}(w_{T+1} + U_{T+1}^0) \qquad (t = 1, \ldots, T+1),$$

where $v(p_t)$ is continuous, strictly increasing and strictly convex for $p_t < p^+$, with $v(0) = 0$ and $\lim_{p_t \to p^+} v(p_t) = \infty$, for p^+ some strictly positive constant.

Additive separability in income and effort is standard in models of moral hazard applied to employment so, in this respect, Assumption 2 is no more restrictive than most of the literature. Linearity in income is a convenient simplification and, with no uncertainty in the model, the degree of risk aversion plays no role. The constant p^+ is an upper bound on the worker's performance. Since the worker gets zero utility in any period of unemployment, it follows that $U_t^0 \geqslant \delta U_{t+1}^0$. U_t^0, Π_t^0 and the functions $y(\cdot)$, $v(\cdot)$, $U_t(\cdot)$, and $\Pi_t(\cdot)$, all $t \geqslant 1$, are common knowledge. Both U_t^0 and Π_t^0 can be functions of any verifiable information but, for simplicity, we do not make this explicit in the notation.

Since the worker's utility is linear in income, the total economic surplus (that is, the sum of the increases in the lifetime utility of the worker and discounted profits of the firm) from the firm employing the worker can be defined independently of the wage paid. The surplus from period t, $t = 1, \ldots, T+1$, on is given by

$$(2.4) \qquad \begin{aligned} S_t(x) &\equiv U_t(x) + \Pi_t(x) - U_t^0 - \Pi_t^0 \\ &= \sum_{\theta=t}^{T} \left[y(p_\theta) - v(p_\theta) \right] \delta^{\theta - t} + \delta^{T-t+1}\left(U_{T+1}^0 + \Pi_{T+1}^0\right) \\ &\quad - \left(U_t^0 + \Pi_t^0\right). \end{aligned}$$

By Assumptions 1 and 2, there exists a unique performance level p^0, defined by

$$(2.5) \qquad p^0 \equiv \underset{p_t \geqslant 0}{\operatorname{argmax}} \left[y(p_t) - v(p_t) \right],$$

that, maintained over a contract of infinite length, maximizes this surplus. We concern ourselves only with the case in which the maximized value of the surplus is nonnegative since, otherwise, the worker would never be employed even if performance were verifiable. If p_t were verifiable, one would then expect the market to establish a Pareto efficient contract that would ensure $p_t = p^0$, for all t, and allocate the surplus so that neither the firm nor the worker was worse off as a result of entering the contract. This provides a useful reference point.

3. THE STRUCTURE OF IMPLICIT CONTRACTS

We model self-enforcing implicit contracts as equilibrium strategies in a dynamic game between the firm and the worker. The rules of the game consist of the explicit part of the contract, enforced by the legal system, which specifies payments conditional on events that are verifiable information. In the present model, the only verifiable pieces of information are the previous employment and wage histories. Whether, in the event of a separation, the worker quit or was fired is not verifiable (though it will obviously be known to the worker and the firm). The reason for this, the practical importance of which is emphasized by Carmichael (1983), is that a firm wanting to fire a worker can make the work environment so intolerable as to induce a quit and a worker wanting to quit can behave so badly as to induce a fire. Hence, any severance payments specified in the explicit part of the contract cannot be conditioned on who initiates the separation.

Payments specified in the explicit part of the contract are denoted $w_t^0 \in (-\infty, \infty)$. The sign and the range of these payments are unrestricted so that an entry fee, or bond posted on entry, is equivalent to $w_1^0 < 0$. Bonds posted at later dates are equivalent to negative severance payments. All such payments are assumed to be enforceable so nothing that follows results from limitations on liability or wealth.

The structure of the game is as follows. In each period $t \geqslant 1$ the worker decides whether to offer to work for the firm ($q_t = 1$) or to quit ($q_t = 0$). Simultaneously, the firm decides whether to hire ($f_t = 1$) or fire ($f_t = 0$) the worker. If $q_t f_t = 1$, the worker sets the level of performance, $p_t \in [0, p^+)$. Then, having observed p_t, the firm pays w_t^0 and, if it wishes, a bonus $w_t^1 \geqslant 0$, giving a total wage of $w_t = w_t^0 + w_t^1$. If $q_t f_t = 0$, the contract is terminated, payments conditional on termination are made, and the worker and the firm expect subsequently to receive the default payoffs U_t^0 and Π_t^0 respectively. Clearly one option for them is to renegotiate another contract with each other, in which case U_t^0 and Π_t^0 are the respective payoffs they expect to receive from renegotiation. Whether or not they renegotiate will depend on the other opportunities available in the market, so we defer discussion of this until we consider market equilibrium in Part II. For the moment, we simply note that both possibilities are covered by the formulation.

Formally, each period t is divided into 3 subperiods, t^0 in which the worker and firm make their simultaneous quit and hire decisions, t^1 in which the worker decides the performance level, and t^2 in which the firm decides the bonus (if any) and pays the wage. Let a_t denote the outcome of these decisions, that is, $a_t = \{q_t, f_t, p_t, w_t^0, w_t^1\}$. The histories that are common knowledge to the worker and the firm are given by

$$h(t^0) = (a_1, a_2, \dots, a_{t-1}), \quad \text{where} \quad h(1^0) = \varnothing,$$
$$h(t^1) = h(t^0) \cup \{q_t, f_t\},$$
$$h(t^2) = h(t^1) \cup \{p_t\}.$$

The histories that are verifiable are given by

$$h^v(t) = \{q_1 f_1, w_1^0, w_1^1, q_2 f_2, w_2^0, w_2^1, \dots, w_{t-1}^0, w_{t-1}^1, q_t f_t\}.$$

We denote by $H(t^i)$, $i = 0, 1, 2$, the set of all possible histories and by $H^v(t)$ the set of all possible verifiable histories.

A strategy $s = (s_1, s_2, \dots)$ for the worker consists of a pair of decision rules specifying q_t and p_t for every information set and all t. Hence, s_t is defined by Q_{t^0}: $H(t^0) \to \{0,1\}$ and P_{t^1}: $H(t^1) \to [0, p^+)$, so that $q_t = Q_{t^0}[h(t^0)]$ and $p_t = P_{t^1}[h(t^1)]$. A strategy $\sigma = (\sigma_1, \sigma_2, \dots)$ for the firm consists of a pair of decision rules specifying f_t and w_t^1 for every information set and all t. Hence σ_t is defined by F_{t^0}: $H(t^0) \to \{0,1\}$ and W_{t^2}: $H(t^2) \to [0, \infty)$, so that $f_t = F_{t^0}[h(t^0)]$ and $w_t^1 = W_{t^2}[h(t^2)]$. It is convenient also to express the explicit part of a contract as a rule $c = \{c_1, c_2, \dots\}$ conditioned on the verifiable history. Then c_t is defined by W_t^0: $H^v(t) \to (-\infty, \infty)$, so that $w_t^0 = W_t^0[h^v(t)]$. For the worker and the firm to be free not to enter a contract implies $W_1^0\{(0)\} = 0$, that is, no payment is required if a contract is not started.

In this notation, c is the explicit part of a contract, (s, σ) the implicit part. We call (s, σ, c) a *contract* and let Ω denote the set of all possible contracts. Each contract results in an allocation x, as defined in (2.1). Let $\gamma[s, \sigma, c|h(t^i)]$, $i = 0, 1, 2$, be the allocation defined by $h(t^i)$ up to t^i and generated by the contract thereafter. The payoffs to the worker and the firm given $h(t^i)$ are then $U_t\{\gamma[s, \sigma, c|h(t^i)]\}$ and $\Pi_t\{\gamma[s, \sigma, c|h(t^i)]\}$, respectively.

DEFINITION: Let $(s^*, \sigma^*, c^*) \in \Omega$. The strategy pair (s^*, σ^*) is an *equilibrium* of the game induced by the explicit contract c^* if, for every $t \geqslant 1$ and $i = 0, 1, 2$, it satisfies the following criteria: (i) (backwards induction)

$$U_t\{\gamma[s^*, \sigma^*, c^*|h(t^i)]\} \geqslant U_t\{\gamma[s, \sigma^*, c^*|h(t^i)]\},$$
$$\text{for all } h(t^i) \in H(t^i) \text{ and } (s, \sigma^*, c^*) \in \Omega,$$

$$\Pi_t\{\gamma[s^*, \sigma^*, c^*|h(t^i)]\} \geqslant \Pi_t\{\gamma[s^*, \sigma, c^*|h(t^i)]\},$$
$$\text{for all } h(t^i) \in H(t^i) \text{ and } (s^*, \sigma, c^*) \in \Omega;$$

(ii) (admissibility) there does not exist an s such that

$$U_t\{\gamma[s, \sigma, c^*|h(t^i)]\} \geqslant U_t\{\gamma[s^*, \sigma, c^*|h(t^i)]\},$$
$$\text{for all } h(t^i) \in H(t^i) \text{ and } (s, \sigma, c^*) \in \Omega,$$
$$\text{with strict inequality holding for some } \sigma \text{ and } h(t^i),$$

and there does not exist a σ such that

$$\Pi_t\{\gamma[s, \sigma, c^*|h(t^i)]\} \geqslant \Pi_t\{\gamma[s, \sigma^*, c^*|h(t^i)]\},$$
$$\text{for all } h(t^i) \in H(t^i) \text{ and } (s, \sigma, c^*) \in \Omega,$$
$$\text{with strict inequality holding for some } s \text{ and } h(t^i).$$

The criterion of backwards induction corresponds to the standard requirement of subgame perfection. To this, we have added the admissibility requirement that the equilibrium strategies are not weakly dominated in any subgame. Kohlberg and Mertens (1986) have argued strongly that admissibility is central to the notion of strategic stability. It also has the direct behavioral interpretation that it is unreasonable for an agent to adopt a strategy when there exists another strategy that is no worse in any outcome of that subgame and is strictly better in some outcome. The relationship of these two criteria to Selten's (1975) definition of perfect equilibrium is discussed by Kohlberg and Mertens (1986) and by van Damme (1983). Following Aumann's (1974) insight of viewing equilibrium points as self-enforcing agreements, we use the following definition.

DEFINITION: A contract $(s^*, \sigma^*, c^*) \in \Omega$ is *self-enforcing* if (s^*, σ^*) is an equilibrium of the game induced by the explicit contract c^*.

The issue of concern here is which allocations can be implemented by self-enforcing contracts. In the next section, we characterize that set of allocations.

4. SELF-ENFORCING IMPLICIT CONTRACTS

Our concern here is with self-enforcing contracts that allow employment to take place. Let $\Omega^*(\Pi^0, U^0)$ denote the set of such contracts for the given sequences of default payoffs $\Pi^0 \equiv (\Pi_1^0, \Pi_2^0, \dots)$ and $U^0 \equiv (U_1^0, U_2^0, \dots)$ and let

$$(4.1) \qquad X^*(\Pi^0, U^0) = \{x \in X \mid x = \gamma(s, \sigma, c) \text{ for some } (s, \sigma, c) \in \Omega^*(\Pi^0, U^0)\}.$$

This is the set of allocations that are implemented by self-enforcing contracts.

PROPOSITION 1: *Under Assumptions* 1–2, *an allocation* $x \in X^*(\Pi^0, U^0)$ *if and only if* T *is infinite and*

$$(4.2) \qquad U_1(x) - U_1^0 \geqslant 0,$$

$$(4.3) \qquad \Pi_1(x) - \Pi_1^0 \geqslant 0,$$

$$(4.4) \qquad \delta\left[U_{t+1}(x) + \Pi_{t+1}(x) - U_{t+1}^0 - \Pi_{t+1}^0\right] \geqslant v(p_t), \quad \text{all} \quad t \geqslant 1.$$

PROOF: (a) *Necessity*. Let $x^\dagger = \{(w_t^\dagger, p_t^\dagger)_{t=1}^{T+1}\} \in X^*(\Pi^0, U^0)$ and suppose $x^\dagger = \gamma(s^\dagger, \sigma^\dagger, c^\dagger)$, where $(s^\dagger, \sigma^\dagger, c^\dagger) \in \Omega^*(\Pi^0, U^0)$. Clearly, (4.2) and (4.3) must hold, otherwise either the worker or the firm would ensure $q_1 f_1 = 0$ and no employment would take place. Moreover, since $(s^\dagger, \sigma^\dagger, c^\dagger) \in \Omega^*(\Pi^0, U^0)$, $U_t(x)$ for each t must be at least as great as if the worker followed any other strategy, in particular, setting $p_t = 0$ and $q_{t+1} = 0$ for some t. Similarly, $\Pi_t(x)$ for each t must be at least as great as if the firm followed any other strategy, in particular, paying the worker as if $p_t = 0$ even if this were not the case and setting $f_{t+1} = 0$ for some t. This gives the following inequalities for all $t = 1, \dots, T$:

$$(4.5) \qquad \begin{aligned} U_t(x^\dagger) \geqslant\ & W_{t^2}\left[h(t^1) \cup \{p_t = 0\}\right] \\ & + \delta\Big\{U_{t+1}^0 + W_{t+1}^0\big[h^v(t) \cup W_t^0(h^v(t)) \\ & \quad \cup W_{t^2}\big(h(t^1) \cup \{p_t = 0\}\big) \cup \{q_{t+1} f_{t+1} = 0\}\big]\Big\}, \end{aligned}$$

$$(4.6) \qquad \begin{aligned} \Pi_t(x^\dagger) \geqslant\ & y(p_t^\dagger) - W_{t^2}\left[h(t^1) \cup \{p_t = 0\}\right] \\ & + \delta\Big\{\Pi_{t+1}^0 - W_{t+1}^0\big[h^v(t) \cup W_t^0(h^v(t)) \\ & \quad \cup W_{t^2}\big(h(t^1) \cup \{p_t = 0\}\big) \cup \{q_{t+1} f_{t+1} = 0\}\big]\Big\}. \end{aligned}$$

Addition of these expressions gives

$$U_t(x^\dagger) + \Pi_t(x^\dagger) \geqslant y(p_t^\dagger) + \delta(U_{t+1}^0 + \Pi_{t+1}^0), \qquad (t = 1, \dots, T),$$

or, in view of the definitions of $U_t(x)$ and $\Pi_t(x)$ in (2.2) and (2.3),

$$\begin{aligned} & y(p_t^\dagger) - v(p_t^\dagger) + \delta\left[U_{t+1}(x^\dagger) + \Pi_{t+1}(x^\dagger)\right] \\ & \qquad \geqslant y(p_t^\dagger) + \delta\left[U_{t+1}^0 + \Pi_{t+1}^0\right], \qquad (t = 1, \dots, T). \end{aligned}$$

In view of Assumptions 1–2, this requires T infinite and thus (4.4) follows.

(b) *Sufficiency*. Consider $x^\dagger = \{(w_t^\dagger, p_t^\dagger)_{t=1}^\infty\} \in X$ satisfying (4.2)–(4.4). We construct a contract $(s^\dagger, \sigma^\dagger, c^\dagger) \in \Omega^*(\Pi^0, U^0)$ that implements $x^\dagger$. For given sequences $\{w_t^m\}_{t=1}^{\infty}$ and $\{w_t^s\}_{t=1}^{\infty}$ (which will be the legally required minimum payment at t if the contract has not been terminated by t and the legally required severance pay if it is terminated at t, respectively), define $(s^{\dagger}, \sigma^{\dagger}, c^{\dagger})$ by

$$(4.7) \qquad s^{\dagger}: Q_{t^0}^{\dagger}[h(t^0)] = \begin{cases} 1, & \text{if } t=1, \text{ or if } p_\theta \geqslant p_\theta^{\dagger} \text{ and } w_\theta \geqslant w_\theta^{\dagger}, 1 \leqslant \theta < t, \\ & \quad \text{or if } w_t^s + U_t^0 \leqslant w_t^m + \delta(w_{t+1}^s + U_{t+1}^0), \ t \geqslant 2; \\ 0, & \text{otherwise;} \end{cases}$$

$$P_{t^1}^{\dagger}[h(t^1)] = \begin{cases} p_t^{\dagger}, & \text{if } t=1, \text{ or if } p_\theta \geqslant p_\theta^{\dagger}, w_\theta \geqslant w_\theta^{\dagger} \text{ and} \\ & \quad q_\theta f_\theta = 1, 1 \leqslant \theta < t; \\ 0, & \text{otherwise;} \end{cases}$$

$$(4.8) \qquad \sigma^{\dagger}: F_{t^0}^{\dagger}[h(t^0)] = \begin{cases} 1, & \text{if } t=1, \text{ or if } p_\theta \geqslant p_\theta^{\dagger} \text{ and } w_\theta \geqslant w_\theta^{\dagger}, 1 \leqslant \theta < t, \\ & \quad \text{or if } -w_t^s + \Pi_t^0 \leqslant -w_t^m - k \\ & \quad +\delta(-w_{t+1}^s + \Pi_{t+1}^0), \ t \geqslant 2; \\ 0, & \text{otherwise;} \end{cases}$$

$$W_{t^2}^{\dagger}[h(t^1)] = \begin{cases} w_{t_2}^{\dagger}, & \text{if } p_\theta \geqslant p_\theta^{\dagger}, w_\theta \geqslant w_\theta^{\dagger} \text{ and } q_\theta f_\theta = 1, \\ & \quad 1 \leqslant \theta \leqslant t, \ t \geqslant 1; \\ 0, & \text{otherwise;} \end{cases}$$

$$(4.9) \qquad c^{\dagger}: W_t^{0\dagger}[h^v(t)] = \begin{cases} w_t^m, & \text{if } q_\theta f_\theta = 1, 1 \leqslant \theta \leqslant t; \\ w_t^s, & \text{if } q_\theta f_\theta = 1 \text{ for } 1 \leqslant \theta < t \text{ and } q_t f_t = 0; \\ 0, & \text{otherwise.} \end{cases}$$

We now show that $\{w_t^m\}_{t=1}^{\infty}$ and $\{w_t^s\}_{t=1}^{\infty}$ can be defined to ensure that, at times t^1, t^2 and t^0, the contract thus defined is self-enforcing, that is, $s^{\dagger}$ and $\sigma^{\dagger}$ are best responses given $c^{\dagger}$ and not weakly dominated.

We first show that, in subgames with $p_{t-1} < p_{t-1}^{\dagger}$ or $w_{t-1} < w_{t-1}^{\dagger}$, it is always a best, and not weakly dominated, response for either the worker or the firm to terminate the contract at t^0. Under $\sigma^{\dagger}$, the firm will do so at $\theta \geqslant t$ unless

$$(4.10) \qquad -w_\theta^s + \Pi_\theta^0 \leqslant -w_\theta^m - k + \delta(-w_{\theta+1}^s + \Pi_{\theta+1}^0).$$

But, since $\Pi_\theta^0 \geqslant \delta \Pi_{\theta+1}^0$, $U_\theta^0 \geqslant \delta U_{\theta+1}^0$ and $k > 0$, then (4.10) satisfied implies

$$(4.11) \qquad w_\theta^s + U_\theta^0 > w_\theta^m + \delta(w_{\theta+1}^s + U_{\theta+1}^0).$$

The lefthand side of (4.11) is the remaining lifetime utility of the worker from quitting at θ^0. Since, under $\sigma^{\dagger}$, $w_\theta = w_\theta^m$, all $\theta \geqslant t-1$, in all subgames with $p_{t-1} < p_{t-1}^{\dagger}$ or $w_{t-1} < w_{t-1}^{\dagger}$, the righthand side is the remaining lifetime utility of the worker at θ^0 from not quitting at θ^0 if a separation would occur at $\theta+1$. It therefore follows that, if when following $\sigma^{\dagger}$ the firm is not going to fire the worker at θ but is going to do so at $\theta+1$, it is a best, and not weakly dominated, response for the worker to quit at θ. But then, moving the argument back one period, for the firm not to have fired the worker at $\theta-1$, it follows that it would have been a best, and not weakly dominated, response for the worker to quit at $\theta - 1$. Sequential application of this argument ensures that, if the firm is ever going to fire the worker and if (4.10) is satisfied for $\theta = t$, then it is a best, and not weakly dominated, response for the worker to quit at t. Moreover, for the firm never to fire the worker under $\sigma^\dagger$ requires that (4.10) is satisfied for all $\theta \geqslant t$. This implies

$$-w_t^s + \Pi_t^0 \leqslant -\sum_{\theta=t}^{\infty} \delta^{\theta-t}(w_\theta^m + k),$$

which, since $\Pi_t^0 \geqslant 0$, $U_t^0 \geqslant 0$ and $k < 0$, implies

$$w_t^s + U_t^0 > \sum_{\theta=t}^{\infty} \delta^{\theta-t} w_\theta^m,$$

so that it is then also a best, and not weakly dominated, response for the worker to quit at t rather than let the contract never be terminated. Hence, whenever (4.10) holds for $\theta = t$, it is a best, and not weakly dominated, response for the worker to quit at t in subgames with $p_{t-1} < p_{t-1}^\dagger$ or $w_{t-1} < w_{t-1}^\dagger$. Moreover, in those subgames the worker will quit at $\theta \geqslant t$ under $s^\dagger$ unless

$$w_\theta^s + U_\theta^0 \leqslant w_\theta^m + \delta\left(w_{\theta+1}^s + U_{\theta+1}^0\right) \tag{4.12}$$

and, by a precisely analogous argument to that above, it is a best, and not weakly dominated, response for the firm to fire the worker at t if (4.12) holds for $\theta = t$. But, since (4.10) implies (4.11), both (4.10) and (4.12) cannot hold for $\theta = t$. Hence, for subgames at t with $p_{t-1} < p_{t-1}^\dagger$ or $w_{t-1} < w_{t-1}^\dagger$, it is always a best, and not weakly dominated, response for either the worker or the firm to terminate the contract at t^0.

Now consider t^1, $t \geqslant 1$. Since the contract will be terminated at $t+1$ if the worker performs below $p_t^\dagger$ whatever the previous history, performing such that $0 < p_t < p_t^\dagger$ is strictly suboptimal. It will also be terminated in any subgame in which $p_\theta < p_\theta^\dagger$ or $w_\theta < w_\theta^\dagger$ for any $\theta < t$, so it is then clearly suboptimal to set $p_t > 0$ and $s^\dagger$ is a best, and not weakly dominated, response in such subgames. Moreover, given $\sigma^\dagger$, it is strictly suboptimal, whatever the history, to set $p_t > p_t^\dagger$ in any subgame as this involves additional disutility of effort for no additional income. Thus, in a subgame with $p_\theta \geqslant p_\theta^\dagger$, $w_\theta \geqslant w_\theta^\dagger$ and $q_\theta f_\theta = 1$ for all $\theta < t$, $s^\dagger$ is a best response at t^1 for all $t \geqslant 1$ if

$$U_t(x^\dagger) \geqslant w_t^m + \delta\left(U_{t+1}^0 + w_{t+1}^s\right), \quad \text{for} \quad t \geqslant 1. \tag{4.13}$$

Moreover, even if this holds with equality, this part of $s^\dagger$ is not weakly dominated since there exist nonequilibrium strategies for the firm (for example, paying $w_t > w_t^\dagger$) that make $p_t = p_t^\dagger$ strictly better than $p_t = 0$.

By a similar argument, strategy $\sigma^\dagger$ is a best, and not weakly dominated, response to $s^\dagger$ at time t^2, $t \geqslant 1$, if

$$\Pi_t(x^\dagger) \geqslant y(p_t^\dagger) - w_t^m + \delta\left(\Pi_{t+1}^0 - w_{t+1}^s\right), \quad \text{for } t \geqslant 1. \tag{4.14}$$

To ensure that the firm pays at least the minimum wage required by $c^\dagger$ requires

(4.15) $\quad w_t^\dagger \geqslant w_t^m, \quad \text{for} \quad t \geqslant 1.$

At time t^0, $t \geqslant 2$, both agents must wish to continue the contract in subgames with $p_{t-1} \geqslant p_{t-1}^\dagger$ and $w_{t-1} \geqslant w_{t-1}^\dagger$ and, since if it is continued we know from the above that the best responses are $p_t = p_t^\dagger$ and $w_t = w_t^\dagger$, sufficient conditions for both wanting to do so are

(4.16) $\quad U_t(x^\dagger) \geqslant U_t^0 + w_t^s, \quad \text{for } t \geqslant 2;$

(4.17) $\quad \Pi_t(x^\dagger) \geqslant \Pi_t^0 - w_t^s, \quad \text{for } t \geqslant 2.$

Again, none of these parts of the strategies are weakly dominated. Moreover, in subgames with $p_{t-1} < p_{t-1}^\dagger$ or $w_{t-1} < w_{t-1}^\dagger$, it is clear that the worker cannot lose by not quitting if (4.12) holds for $\theta = t$ since the righthand side can be guaranteed by quitting at $t+1$ and it is not weakly dominated since the worker would do even better by staying on if the firm did not fire and were to pay $w_t > w_t^m$. Similarly, in these subgames it is a best, and not weakly dominated, response for the firm not to fire if (4.10) holds for $\theta = t$.

This establishes that (4.13)–(4.17) are sufficient for $s^\dagger$ and $\sigma^\dagger$ to be best, and not weakly dominated, responses given $c^\dagger$ in all subgames after t^0 for $t = 1$. But, if one selects $w_t^s = U_t(x^\dagger) - U_t^0$ and $w_t^m = w_t^\dagger - v(p_t^\dagger)$, then (4.4) is sufficient to ensure that (4.13)–(4.17) hold by the following argument. (4.16) and (4.15) follow directly from the way w_t^s and w_t^m are defined and $v(p_t) \geqslant 0$. Moving (4.4) back one period and replacing $U_t(x^\dagger) - U_t^0$ by w_t^s, one gets that (4.17) is satisfied. Moreover, using the definitions

$$U_t(x^\dagger) = w_t^\dagger - v(p_t^\dagger) + \delta U_{t+1}(x^\dagger),$$

$$\Pi_t(x^\dagger) = y(p_t^\dagger) - w_t^\dagger + \delta \Pi_{t+1}(x^\dagger),$$

one gets that (4.13) and (4.14) are satisfied.

Finally, given that $s^\dagger$ and $\sigma^\dagger$ are best responses in all subsequent subgames, (4.2) and (4.3) ensure that, at time 1^0, entering the contract is a best response for both worker and firm. This is not weakly dominated even when one of these holds with equality since there are clearly nonequilibrium strategies of the other party that make entering the contract strictly better than not doing so. *Q.E.D.*

Proposition 1 characterizes the set of allocations that can be implemented by self-enforcing contracts. Note two technical points about the proof. First, the necessity part does not use the requirement that equilibrium strategies be admissible, so dropping that requirement would not extend the set of allocations implemented by self-enforcing contracts. Second, the proof that the quitting strategy of the worker in (4.7) and the firing strategy of the firm in (4.8) are admissible best responses does not in fact depend on those decisions being simultaneous as we have assumed. Thus the set of allocations implementable by self-enforcing contracts is independent of whether these decisions are made simultaneously or sequentially, in either order.

The formal proof of Proposition 1 is complicated by the need to deal with all subgames but the underlying idea is straightforward. Conditions (4.2) and (4.3) are obvious requirements since, unless they are satisfied, either the worker or the firm will be better off out of the contract. Condition (4.4) is necessary for the following reason. Basically, the verifiable information permits only two types of payment to be legally enforceable, a minimum wage w_t^m if the contract continues and a severance payment w_t^s (of either sign) if it terminates. Suppose the implicit part of the contract specifies a bonus of $w_t^1 \equiv w_t - w_t^m \geqslant 0$ in period t if performance is p_t. Then performing at p_t is worthwhile for the worker only if the bonus plus the gain, discounted back to t, from having the contract continue is at least as great as the disutility of effort plus the (discounted) severance payment, that is, if $w_t^1 + \delta[U_{t+1}(x) - U_{t+1}^0] \geqslant v(p_t) + \delta w_{t+1}^s$. Paying the bonus is worthwhile for the firm only if the gain, discounted back to t, from having the contract continue is at least as great as the bonus minus the (discounted) severance payment, that is, if $\delta[\Pi_{t+1} - \Pi_{t+1}^0] \geqslant w_t^1 - \delta w_{t+1}^s$. Adding these two conditions, we get (4.4). Moreover, if (4.4) is satisfied, one can clearly always find a value of $w_t^1 - \delta w_{t+1}^s$ with $w_t^1 \geqslant 0$ that ensures that these two conditions are satisfied, which is the basis of the proof of sufficiency. Indeed, this can be done with any $w_t^1 > 0$ so that, contrary to what has been assumed in the literature, implicit contracts in which the wage depends on performance (that is, bonus or piece-rate contracts) can be made self-enforcing even when performance is not verifiable.

For a nonterminating contract, (4.4) can, in view of (2.2) and (2.3), be written

$$\delta\left\{\sum_{\theta=t+1}^{\infty}[y(p_\theta) - v(p_\theta)]\delta^{\theta-t+1} - \left(U_{t+1}^0 + \Pi_{t+1}^0\right)\right\} \geqslant v(p_t), \quad \text{all} \quad t \geqslant 1.$$

This depends only on the sequence of performance levels in an allocation, not on the sequence of wage payments. It follows that only the present discounted value of wage payments, not the time path, is important for the implementability of an allocation. It also follows that, to implement a nontrivial allocation, there must always exist a strictly positive surplus from continuing the contract once it has started, that is, for $t \geqslant 2$, since the term in braces on the lefthand side of this is just the surplus at $t+1$ as defined in (2.4).

For some values of U_1^0 and Π_1^0 there may also be a strictly positive surplus at $t = 1$. That surplus can be divided between the firm and the worker in any proportion provided an appropriate contract is chosen. Contracts can be found which give the worker all the surplus, the firm all the surplus, or any division between them. To see this, note that the expression for the surplus in (2.4) for $t = 1$ is the sum of the lefthand sides of (4.2) and (4.3), the worker's part being given by the lefthand side of (4.2) and the firm's by the lefthand side of (4.3). But, from (2.4), the surplus is independent of w_1 so that, without affecting (4.4), w_1 can be altered to divide the surplus in any proportions.

Proposition 1 allows the default payoffs to follow any paths. In the present stationary model, however, one would expect the default utility and profit to be constant through time. Then one can show that any efficient incentive compatible allocation has performance constant through time and, for any such efficient allocation, there is an equally efficient allocation with the wage constant through time. It is also straightforward to derive necessary and sufficient conditions for $X^*(\Pi^0, U^0)$ to be nonempty.

DEFINITION: An allocation $x \in X^*(\Pi^0, U^0)$ is *efficient* if there does not exist an $x' \in X^*(\Pi^0, U^0)$ such that $U_1(x') \geqslant U_1(x)$, $\Pi_1(x') \geqslant \Pi_1(x)$ and either $U_1(x') > U_1(x)$ or $\Pi_1(x') > \Pi_1(x)$.

ASSUMPTION 3: $U_t^0 = u^0/(1-\delta)$ *and* $\Pi_t^0 = \pi^0/(1-\delta)$, *all* $t \geqslant 1$, *for some constants* u^0, $\pi^0 \geqslant 0$.

PROPOSITION 2: *Under Assumptions* 1–3,
(a) *there exists an* $x \in X^*(\Pi^0, U^0)$ *if and only if*

$$(4.18) \qquad \max_{p_t \geqslant 0} \left[y(p_t) - v(p_t)/\delta \right] \geqslant u^0 + \pi^0;$$

(b) *if* (4.18) *is satisfied*, $x^* \equiv \{(w_t^*, p_t^*)_{t=1}^{\infty}\} \in X^*(\Pi^0, U^0)$ *is efficient only if* $p_t^* = p^\dagger$, *all* t, *where* $p^\dagger$ *is uniquely defined by*

$$(4.19) \qquad p^\dagger = \arg\max_{p_t \geqslant 0} y(p_t) - v(p_t),$$

$$\text{subject to} \quad y(p_t) - v(p_t)/\delta - (u^0 + \pi^0) \geqslant 0;$$

(c) *if* (4.18) *is satisfied and if* $x^* \equiv \{(w_t^*, p_t^*)_{t=1}^{\infty}\} \in X^*(\Pi^0, U^0)$ *is efficient, there exists a* $w^\dagger$ *such that the stationary allocation* $x^\dagger \equiv \{(w^\dagger, p^\dagger), (w^\dagger, p^\dagger), \dots\} \in X^*(\Pi^0, U^0)$ *and has* $U_1(x^\dagger) \geqslant U_1(x^*)$ *and* $\Pi_1(x^\dagger) \geqslant \Pi_1(x^*)$.

PROOF: We start with part (b). Note from (2.4) that choosing x to maximize the surplus $S(x)$ is equivalent to choosing x to maximize $U_1(x) + \Pi_1(x)$. But, also from (2.4), $S(x)$ depends only on the sequence of performance levels and is independent of the sequence of wage payments. Thus any distribution of the surplus can be achieved independently of performance by suitable adjustment of wage payments and any efficient allocation must have a performance sequence that maximizes the surplus. Let $\Xi = \{(p_1, p_2, \dots) | p_t \in [0, p^+)$, all $t \geqslant 1\}$ denote the set of all possible sequences of performance levels. For $\xi \in \Xi$, let $S_t^*(\xi)$ denote the surplus defined by (2.4) but written as a function of ξ. Define $G_t(\xi) = S_t^*(\xi) - v(p_{t-1})/\delta$ for $t \geqslant 2$. Note that $G_t(\xi) \geqslant 0$ is equivalent to condition (4.4). The proof proceeds by first showing that $x^* \in X^*(\Pi^0, U^0)$ implies that there is a stationary sequence of performance levels $\xi = (p, p, \dots)$ satisfying $G_t(\xi) \geqslant 0$ for every $t \geqslant 2$. Then we show that the stationary sequence of performance levels $p^\dagger$ defined in (4.19) uniquely solves

$$(4.20) \qquad \max_{\xi \in \Xi} S_1^*(\xi) \text{ subject to } G_t(\xi) \geqslant 0, \text{ for } t \geqslant 2.$$

Let $\xi^* \equiv (p_1^*, p_2^*, \dots)$ be the sequence of performance levels corresponding to the allocation x^*. Then, from (4.4) of Proposition 1, $G_t(\xi^*) \geqslant 0$ for $t \geqslant 2$. Let Λ, defined by $\Lambda(\xi) = (p_2, p_3, \dots)$, denote the shift operator on Ξ. Then, from Assumption 3, it is clear that $G_t[\Lambda^n(\xi^*)] \geqslant 0$ for every $t \geqslant 2$ and $n \geqslant 0$. Moreover, by Assumptions 1 and 2, the functions $G_t(\xi)$ are strictly concave for all $\xi \in \Xi$; hence $G_t[N^{-1} \cdot \Sigma_{n=0}^{N} \Lambda^n(\xi^*)] \geqslant 0$ for every $t \geqslant 2$ and $N \geqslant 0$. Let $p' = \limsup_{N\to\infty} N^{-1} \cdot \Sigma_{t=1}^{N} p_t^*$ and let $\xi' = (p', p', \dots)$. For every $t \geqslant 2$, the functions $G_t(\cdot)$ are continuous with respect to the l_∞ norm and, hence

$$\limsup_{N\to\infty} G_t\left[N^{-1} \cdot \sum_{n=0}^{N} \Lambda^n(\xi^*)\right] = G_t(\xi') \geqslant 0.$$

Thus the program in (4.19) is feasible and, by Assumptions 1 and 2, the performance level $p^\dagger$ defined by it is unique.

The sequence of performance levels $\xi^\dagger = (p^\dagger, p^\dagger, \dots)$ is clearly feasible for the program (4.20). Moreover, since program (4.20) is strictly concave, it is a straightforward exercise to show that $\xi^\dagger$ uniquely solves it. This establishes part (b). To establish part (c), note that for x^* efficient $p_t^* = p^\dagger$, all t, follows immediately from (b). Moreover, there then exists a $w^\dagger \geqslant 0$ such that

$$(4.21)\quad \begin{aligned} &\left[w^\dagger - v(p^\dagger)/(1-\delta)\right] \geqslant \sum_{t=1}^{\infty} \left[w_t^* - v(p_t^*)\right]\delta^{t-1} = U_1(x^*), \\ &\left[y(p^\dagger) - w^\dagger\right]/(1-\delta) \geqslant \sum_{t=1}^{\infty} \left[y(p_t^*) - w_t^*\right]\delta^{t-1} = \Pi_1(x^*), \end{aligned}$$

so (4.2) and (4.3) are certainly satisfied for $x^\dagger$ and (4.4) is satisfied since $G_t(\xi^\dagger) \geqslant 0$, all $t \geqslant 2$. Thus, the conditions of Proposition 1 are satisfied for the stationary allocation $x^\dagger = \{(w^\dagger, p^\dagger), (w^\dagger, p^\dagger), \dots\}$ and so $x^\dagger \in X^*(\Pi^0, U^0)$. This establishes part (c). The necessity part of (a) holds because (4.4) is equivalent to $G_t(\xi) \geqslant 0$ for all $t \geqslant 2$ and from (b) there exists an x with performance sequence satisfying this only if there exists a p_t satisfying the constraint in (4.19). The sufficiency part of (a) follows from our demonstration that $x^\dagger \in X^*(\Pi^0, U^0)$.
Q.E.D.

Part (a) of this gives a necessary and sufficient condition for existence of a nontrivial allocation that can be implemented by a self-enforcing contract. Parts (b) and (c) characterize efficient contracts. It is not surprising that, in a stationary model, nonstationary allocations cannot improve on stationary ones. That a stationary performance strictly dominates nonstationary ones follows essentially from the strict convexity of $v(\cdot)$. It also has an interesting consequence. With stationary default payoffs and stationary performance, the surplus defined in (2.4) is the same at each date. Since we know from Proposition 1 that there must be a strictly positive surplus from period 2 on, there must then also be a strictly positive surplus in period 1. This means that it is not possible for both the firm and the worker to receive payoffs equal to their default payoffs when agreeing to a contract. One or both of them must be made strictly better off by doing so.

Even with stationary allocations, the surplus required for a contract to be self-enforcing can be divided in any way between the worker and the firm. Since, however, a payment in period 1 different from that in subsequent periods would violate stationarity, the nature of the contract must change with a change in the division of the surplus. The form of the contract is the subject of Proposition 3. For notational convenience, we let $w_t^m(c)$ be the minimum payment required by the explicit part c of a contract that has not been terminated by t and $w_t^s(c)$ the severance payment required by c if it is terminated at t.

PROPOSITION 3: *Let* $x^\dagger = \{(w^\dagger, p^\dagger), (w^\dagger, p^\dagger), \dots\} \in X^*(\Pi^0, U^0)$ *and let* $S^w(x^\dagger) \equiv [U_1(x^\dagger) - U_1^0]/S_1(x^\dagger)$ *denote the share of the surplus that goes to the worker. Then, under Assumptions* 1–3, $(w^\dagger, p^\dagger)$ *satisfies*

$$(4.22)\qquad w^\dagger = y(p^\dagger) - \pi^0 - (1-\delta)S_1(x^\dagger)[1 - S^w(x^\dagger)] = v(p^\dagger) + u^0 + (1-\delta)S_1(x^\dagger)S^w(x^\dagger)$$

and the explicit part of the contract c implements $x^\dagger$ *if and only if*

$$(4.23)\qquad w_t^m(c) \leqslant w^\dagger, \quad \text{all } t,$$

$$(4.24)\qquad y(p^\dagger) - \pi^0 - S_1(x^\dagger)[1 - S^w(x^\dagger)] \leqslant w_t^m(c) + \delta w_{t+1}^s(c) \leqslant u^0 + S_1(x^\dagger)S^w(x^\dagger), \quad \textit{all } t.$$

Moreover, there exists c such that (4.23) *and* (4.24) *are satisfied for every t with* $w_{t+1}^s(c) = 0$, *that is, with no severance pay or bonding.*

PROOF: With $x^\dagger$ stationary, the definition of the shares of the surplus implies

$$S_1(x^\dagger)S^w(x^\dagger) = [w^\dagger - v(p^\dagger) - u^0]/(1-\delta),$$

$$S_1(x^\dagger)[1 - S^w(x^\dagger)] = [y(p^\dagger) - w^\dagger - \pi^0]/(1-\delta),$$

from which (4.22) immediately follows. Necessity of (4.23) follows from the definition of a contract. Necessity of (4.24) follows because, under Assumption 3 the necessary conditions (4.5) and (4.6) can be written

$$[w^\dagger - v(p^\dagger)]/(1-\delta) \geqslant w_t^m(c) + \delta u^0/(1-\delta) + \delta w_{t+1}^s(c), \quad \text{all } t,$$

$$[y(p^\dagger) - w^\dagger]/(1-\delta) \geqslant [y(p^\dagger) - w_t^m(c)] + \delta\pi^0/(1-\delta) - \delta w_{t+1}^s(c), \quad \text{all } t.$$

The inequalities (4.24) follow directly from these and the definitions of $S_1(x)$ in (2.4) and $S^w(x)$. Proof of sufficiency of (4.23) and (4.24) follows the sufficiency proof of Proposition 1 with (w_t^m, w_{t+1}^s) satisfying (4.23) and (4.24), all t. Finally (4.22)–(4.24) can always be satisfied with $w_{t+1}^s(c) = 0$, all t, provided

$$y(p^\dagger) - \pi^0 - S_1(x^\dagger)[1 - S^w(x^\dagger)] \leqslant y(p^\dagger) - \pi^0 - (1-\delta)S_1(x^\dagger) \times [1 - S^w(x^\dagger)].$$

But this is necessarily the case since $0 < \delta < 1$. *Q.E.D.*

For stationary allocations, Proposition 3 specifies how the form of contract depends on the division of the surplus and establishes that the use of severance pay and bonding does not extend the set of implementable allocations. As is clear from (2.4), the surplus $S_1(x^\dagger)$ depends only on $p^\dagger$ and not on $w^\dagger$. Hence, for given $p^\dagger$, (4.22)–(4.24) specify completely what the terms of the explicit part of the contract must be if the worker is to receive the share of the surplus $S^w(x^\dagger)$. If $S^w(x^\dagger) = 0$, it follows from the proposition that either $w^\dagger > w_t^m(c)$ or $w_{t+1}^s(c) < 0$, all t, and hence that either the contract has a bonus or piece-rate element or that it has negative severance pay (that is, bonding). The reason is that a worker who gets no surplus gets no long-term benefit from continuing the contract and will, therefore, default unless there is an immediate cost to doing so. If, on the other hand, $S^w(x^\dagger) = 1$, then $w_t^m(c) + \delta w_{t+1}^s(c) \geqslant w^\dagger$ and, since $w_t^m(c) \leqslant w^\dagger$ by definition, $w_{t+1}^s(c) \geqslant 0$. In this case, therefore, there cannot be bonding and, if the contract has a bonus or piece-rate element, it must also have strictly positive severance pay. The reason in this case is that, if the worker gets all the surplus, the firm gets no long-term benefit from continuing the contract and would therefore default on it if there were bonding or if the contract were a piece-rate with no severance pay.

Shapiro and Stiglitz (1984) assume that contracts take the form $w_t^m(c) = w^\dagger$ and $w_{t+1}^s(c) = 0$, that is, they are termination contracts with no severance pay or bonding. Since $0 < \delta < 1$, it then follows from (4.23) and the first halves of (4.22) and (4.24) that $S^w(x^\dagger) = 1$, so the worker gets all the surplus. Proposition 3, however, makes clear that this does not have to be the case, even with a stationary allocation, provided piece-rate contracts or bonding are permitted. In that case, there is no necessity for the worker to receive any of the surplus.

The result that any achievable stationary allocation can be implemented by a contract c with $w_{t+1}^s(c) = 0$, that is, without severance pay or bonding, may seem surprising. Bonding has been widely discussed in the literature as a solution to the problem of worker moral hazard. See Eaton and White (1982) and (1983) on this. The economic rationale for the result is, however, straightforward. *Ceteris paribus*, a bond clearly increases the loss to the worker from dismissal and so reduces the incentive to cheat on the implicit contract. However, in the present model, it simultaneously increases by exactly the same amount the gain to the firm from defaulting on the implicit contract by dismissing the worker even if performance has been at the appropriate level. Hence a bond cannot reduce the total surplus required to make a contract self-enforcing. All it can do is allow more of the surplus to go to the firm without inducing the worker to default and that can be done just as well by piece-rate or bonus contract.

That neither severance pay nor bonding is necessary to achieve any implementable allocation has an important consequence. In the absence of these, the only legally enforceable element in a contract is the part of the wage that is independent of the worker's performance. Since this must be paid whatever the performance, it would make no difference to the incentives if it were paid at the time the worker reported for work at the beginning of the period instead of being paid at the end, as we have assumed. But then there is no need for a legally enforceable contract to ensure that the firm pays it because the worker need not work unless paid. Note that it is the absence of severance pay and bonding that is important for this. These (in the way we have formulated them) are payments made only if a contract has terminated but, once a contract has terminated, there is no incentive for them to be paid unless they are legally enforceable. Without severance pay and bonding, however, wage payments under a contract can be scheduled to occur only at times when the firm gains by paying them and continuing the contract, so it is in its interest to pay them even without legal enforceability.

The results of this section are important from a theoretical point of view because they give a complete characterization of all the allocations that can be implemented by any self-enforcing contract. This allows us to discuss the implications of nonverifiability without having to be concerned, as Carmichael (1985) rightly was about Shapiro and Stiglitz (1984), that the conclusions we draw result from the assumption of a specific form of contract. Moreover, in discussing market equilibria in Part II of this paper, it enables us to concern ourselves only with allocations, rather than specific contracts, because we know that there always exists a self-enforcing contract that will implement any allocation satisfying the conditions of Proposition 1.

These results are also important in providing a better understanding of the labor contracts used in practice. Piece-rate contracts, and contracts with an informally agreed bonus, are used in situations in which their legal enforceabiliity is at best doubtful. Severance payments and bonding are not particularly widespread, despite their prominence in earlier literature concerned with moral hazard in employment. And labor markets survived with few written contracts for a very long time despite the fact that the lack of a written contract greatly increases the difficulty of legal enforcement. It does seem, therefore, that the contracts that arise from the present approach are much closer to actual contracts than those that arise from applying standard principal-agent models to employment.

PART II: MARKET EQUILIBRIUM

5. SUBGAME PERFECT EQUILIBRIA

Part I of this paper has characterized the set of wage-performance allocations that can be implemented by implicit contracts between a firm and an employee when the employee's performance is not verifiable in court. There the default payoffs received if the contract was terminated were taken as given. In Part II, we embed our bilateral contract structure into a market setting in which the default payoffs are endogenous. Our purpose is twofold. First, we want to analyze the implications of a competitive labor market for the set of allocations implementable by self-enforcing contracts. Second, we want to compare the characteristics of equilibrium in labor markets in which performance is not verifiable with that in labor markets without this information imperfection.

From the formal point of view, this requires extending the two-player game of Part I into a multiple-player game. The fundamental issues raised by this, however, apply to a wider set of models than that of Part I. The essential characteristic of that model for market equilibrium is that mutually beneficial contracts between two parties are self-enforcing only when at least one party gains from continuing the contract and is deterred from deviating for additional short-term gain by the threat of the other party playing a punishment strategy. This is a characteristic that arises in other repeated games with equilibria sustained by the threat of punishment strategies, for example, the repeated principal-agent models studied in Radner (1985). The existence of a market in which the players can find other partners then has two important consequences. First, a player who deviates can negotiate with another partner to avoid the full force of the punishment inflicted by the other party. But, second, a punishing party can also negotiate with another partner and that opportunity may reduce the cost, and hence increase the credibility, of adopting a punishment strategy. In this respect the fundamentals, though not the details, of the analysis that follows apply to other market game with these characteristics.

The model we use is identical with that of Part I with the following additional assumption that specifies the structure of the economy.

ASSUMPTION 4: *The economy has one produced good, which is not storable, is used as the numeraire, and is produced by workers each working with one production function* $y(p_t)$. *There are* N *such production functions (jobs), indexed by* $n \in \Upsilon \equiv \{1, \ldots, N\}$, *and* L *identical workers, indexed by* $l \in \mathscr{L} \equiv \{1, \ldots, L\}$. *The owners of production functions (firms) do not work and consume all profits.*

There are a number of different ways one might specify the rules of the game that govern the operation of the labor market. We specify one particular set here and discuss later how the results would differ under other specifications. Our description is somewhat informal. A more formal description along the lines of that in Section 3 requires a lot of notation that obscures, rather than clarifies, the basic points.

A labor market opens in each period before production begins. In period 1, all agents in the economy enter this market. In subsequent periods, it is entered only by workers who were unemployed in, or have separated from their employer of, the previous period, and by firms that had an unfilled vacancy in, or that separated from an employee of, the previous period. Let $\mathscr{L}^t$ denote the set of workers in the market in period t and Υ^t the set of jobs that firms have vacant. When the market opens, firms announce a contract $\omega^n \equiv (s^n, \sigma^n, c^n) \in \Omega$ for each vacancy $n \in \Upsilon^t$. Note that Ω contains the trivial contract of not offering employment. The worker with the lowest index in $\mathscr{L}^t$ then chooses an unfilled vacancy $n \in \Upsilon^t$ by agreeing formally to the explicit conditions c^n, or else chooses to remain unemployed. Once this choice has been made, the worker with the next lowest index makes a similar choice from the remaining vacancies. This process continues until there are either no vacancies or no workers left in the labor market. Workers and firms that have formed a match then play the game described in Section 3 under the explicit conditions c^n for each $n \in \Upsilon^t$.

As in Part I, equilibria of this game are required to satisfy the backward induction and admissibility criteria of Kohlberg and Mertens (1986). The formal definition is an obvious extension of that in Section 3 so we do not give details. A market allocation can be denoted by $(x^1, \ldots, x^E) \in X^E$, where $E \leqslant \min\{N, L\}$ is the employment level and $x^i \in X$, $i = 1, \ldots, E$, is the allocation in the ith match. Since all workers are identical, as are all jobs, there is no need to specify which worker and job form the ith match so, without loss of generality, we can take it that job $n = i$ and worker $l = i$ form the ith of the E matches and adopt the convention that $\Pi_t(x^i) = U_t(x^i) = 0$, all $t \geqslant 1$, for $i > E$ so that the payoffs to those not in matches are zero. Recall from (4.1) that $X^*(\Pi^0, U^0)$, characterized in Proposition 1, is the set of allocations for a job that can be implemented by self-enforcing contracts given the sequences of default payoffs $\Pi^0 \equiv (\Pi_1^0, \Pi_2^0, \ldots)$ and $U^0 \equiv (U_1^0, U_2^0, \ldots)$. We let $\varnothing$ denote the null sequence in which each term is zero. The set of market allocations that can be implemented by equilibria are characterized as follows.

PROPOSITION 4: *Under Assumptions* 1, 2, *and* 4, *a market allocation* $(x^1, \ldots, x^E) \in X^E$ *can be implemented by an equilibrium if and only if*

(5.1) $\quad 0 \leqslant E \leqslant \min\{N, L\};$

(5.2) $\quad x^i \in X^*(\varnothing, \varnothing), \quad \text{for } i = 1, \ldots, E.$

PROOF: (a) *Necessity*. Let $(x^1, \ldots, x^E)$ be a market allocation implemented by an equilibrium. Then (5.1) follows immediately from the requirement that jobs and workers form one-to-one matches. Moreover, workers and firms can always achieve the payoff sequence $\varnothing$ by refusing any matches so, in equilibrium, their default payoffs must be at least as great as this and thus (5.2) must hold.

(b) *Sufficiency*. Let $(x^1, \ldots, x^E)$ be a market allocation satisfying (5.1) and (5.2). The proof proceeds by constructing strategies that form an equilibrium and implement this allocation. In particular, the strategies are constructed to yield a zero payoff to all agents when off the equilibrium path. To save on heavy notation, the specification of these strategies is somewhat informal. Let $w_t^m(c)$ denote the minimum payment in period t under the explicit contract c if that contract has not been terminated by t and $w_t^s(c)$ the severance payment required at t if the contract is terminated in period t. Also denote by Υ_l^t the set of jobs still vacant at t after workers with index less than l have chosen either jobs or unemployment.

Strategy of firms for job $n \geqslant E + 1$: If in the labor market in period t, offer a contract c^n with $-w_t^m(c^n) - \delta w_{t+1}^s(c^n) - k \geqslant 0$. In any subgame in period t in which the job is filled, pay $w_t^m(c^n)$ and, if $-w_{t+1}^m(c^n) - \delta w_{t+2}^s(c^n) - k \geqslant -w_{t+1}^s(c^n)$, keep the worker on for period $t+1$; otherwise fire the worker.

Strategy of firms for job $n \in \{1, \ldots, E\}$: In period 1, offer a contract $\omega^n \in \Omega^*(\varnothing, \varnothing)$ for that job that is both self-enforcing and implements x^n (such a contract exists by Proposition 1) and follow the strategy σ^n specified by this contract until it is terminated. In every other subgame, follow the strategy specified for jobs with index $n \geqslant E+1$. (Note that, even if $E = N$, these strategies are well defined.)

Strategy of worker $l \in \mathscr{L}$: Denote by Υ_l^* the set of vacancies available to worker l in period 1 for which the allocation specified by (5.2) is offered under a self-enforcing contract. Formally,

$$\Upsilon_l^* \equiv \{ n | n \in \{1, \ldots, E\} \cap \Upsilon_l^1 \text{ and } \omega^n \in \Omega^*(\varnothing, \varnothing), \text{ is self-enforcing and implements } x^n \}.$$

Denote by n' the job that maximizes $U_1(x^n)$ among $n \in \Upsilon_l^*$. Let $n' = 0$ if no such job exists and, in the event of a tie, choose randomly among the tying jobs. Denote by n'' the job that has the highest value of $w_1^m(c^n) + \delta w_2^s(c^n)$ among $n \in \Upsilon_l^1 - \Upsilon_l^*$. In period 1 choose job n', n'', or no employment according as $U_1(x^{n'})$, $w_1^m(c^{n''}) + \delta w_2^s(c^{n''})$, or 0 is the greater. In the event of a tie, choose n' before n'' and, likewise, n'' before no employment. If n' is chosen in period 1, follow the strategy $s^{n'}$ specified by the implicit part of $\omega^{n'}$ until the contract is terminated. In all other subgames, adopt the following strategy: (i) if in the labor market in period $t \geqslant 2$, choose the job that offers the largest payment $w_t^m(c^n) + \delta w_{t+1}^s(c^n)$ for $n \in \Upsilon_l^t$ if this amount is at least zero, otherwise remain unemployed; (ii) if employed in job n in period $t \geqslant 1$, set performance p_t equal to zero and if $w_{t+1}^s(c^n) > w_{t+1}^m(c^n) + \delta w_{t+2}^s(c^n)$ quit, otherwise stay on.

The workers' strategy is constructed so that, if the contract offered for job $n \in \{1, \ldots, E\}$ is self-enforcing and implements x^n, some worker accepts it and does not subsequently deviate from the equilibrium path for that contract whereas, if the contract is not self-enforcing or does not implement x^n, either no worker accepts it or else a worker accepts it but sets performance at zero. In view of that, the firm receives profits $\Pi_1(x^n) \geqslant 0$ from job $n \in \{1, \ldots, E\}$ by offering ω^n as in the strategy specified above and following σ^n whereas, if it offers a contract that is not self-enforcing or does not implement x^n, it receives a payoff of at most zero, so following that strategy is a best response. Clearly, if $\Pi_1(x^n) > 0$ this is not weakly dominated but this is true even if $\Pi_1(x^n) = 0$ since filling the vacancy with allocation x^n is strictly better for some (nonequilibrium) strategy of workers than any alternative. Given firms' strategies, a worker who chooses job $n \in \{1, \ldots, E\}$ and follows the strategy s^n gets utility $U_1(x^n) \geqslant 0$ so it is clearly a best response for workers to select jobs as described in the above strategy and, by an argument similar to that for firms, this is not weakly dominated. Moreover, if either party deviates from the equilibrium path of ω^n, the strategies of other players ensure that both receive a payoff of zero thereafter. Thus, since $x^n \in X^*(\varnothing, \varnothing)$ and ω^n implements x^n, it is a best, and not weakly dominated, response for neither party to deviate from the equilibrium path of ω^n.

To show that the strategies form an equilibrium of the market game, it remains to show that the strategies for workers and jobs not matched in one of the E initial matches, or whose initial contract is terminated, are best responses and not weakly dominated. By exactly the same argument used in the proof of Proposition 1, in any such subgame it is always a best, and not weakly dominated, response for either the worker or the firm to terminate any contract. Hence, a worker who finds himself or herself employed has an income independent of effort and thus setting performance p_t equal to zero is always the unique best response. In view of this, accepting the vacancy that offers the largest payment $w_t^m(c^n) + \delta w_{t+1}^s(c^n)$ if this amount is at least zero is also always a best response and cannot be weakly dominated. Moreover, by that same argument, it is uniquely optimal for the worker to quit unless $w_{t+1}^s(c^n) \leqslant w_{t+1}^m(c^n) + \delta w_{t+2}^s(c^n)$. If that condition holds, then the worker loses nothing by staying on as he or she can always quit with effect from period $t+2$ and thus receive $w_{t+1}^m(c^n) + \delta w_{t+2}^s(c^n)$ and this is clearly not weakly dominated.

Given the strategy of workers, offering a contract with $-w_t^m - \delta w_{t+1}^s - k \geqslant 0$ is a best, and not weakly dominated, response for the firm with job $n \geqslant E+1$, or $n \in \{1, \ldots, E\}$ whose initial contract is terminated, because any worker accepting a contract with that firm will shirk (but no worker actually accepts such a contract). Equally, if the firm has an employee, the employee will shirk so paying the lowest wage permitted under the contract is a best response and also clearly not weakly dominated. Finally, by an argument analogous to that for workers, it is a best, and not weakly dominated, response for the firm to fire an employed worker unless $-w_{t+1}^m(c^n) - \delta w_{t+2}^s(c^n) - k \geqslant -w_{t+1}^s(c^n)$.

Hence, the strategies specified above form a subgame perfect, not weakly dominated equilibrium that implements the market allocation $(x^1, \ldots, x^E)$.

Q.E.D.

Condition (5.1) of this proposition states that any employment level up to the maximum feasible level of $\min\{N, L\}$ is a possible market equilibrium. Condition (5.2) states that market equilibrium places no restrictions on the allocations beyond those required by Proposition 1 to ensure that they can be implemented by a self-enforcing contract when the default utility and profit levels are zero. This is in the spirit of the "Folk Theorem" results for repeated games (see Fudenberg and Maskin (1986)); taken in conjunction with Proposition 1, (5.2) implies that any self-enforcing combination of firms' profits and workers' utilities that gives all agents payoffs at least as good as if they did not participate in the market is an equilibrium.[3] These equilibria are sustained by punishment strategies in the usual way. If in period 1 a firm offers anything other than a contract that implements the equilibrium allocation for its job, either no worker accepts or

[3] This is different from the Folk Theorem results in Fudenberg and Maskin (1986) in several ways. Particularly important in the present context are: (a) the set of equilibria is characterized for all δ, not just δ close to 1; (b) the equilibria include any or all agents receiving their minimax payoffs; (c) agents are not assumed to observe *all* other agents' past strategies; (d) we have ensured that equilibrium strategies are admissible.

(b) is necessary for equilibria with $E < \min\{N, L\}$. (c) is important because the structure of our model rules out firms observing the performance of workers who are not their employees and of workers observing the performance of other workers. This is clearly important in the context of labor markets.

a worker accepts and shirks. If either firm or worker fails to abide by such a contract, they are punished by the other terminating it. In any other subgame, workers shirk and firms fire any workers they have, so any agent who deviates from the equilibrium allocation is punished by receiving a nonpositive payoff thereafter. Since any agent can guarantee a payoff of zero by not participating in the market, the default payoffs Π_t^0 and U_t^0 are zero in every period and any allocation $x \in X^*(\varnothing, \varnothing)$ can be implemented as an equilibrium.

Both (5.1) and (5.2) are different from what one expects for equilibria in models in which performance is verifiable. For $E < \min\{N, L\}$, there exist bilateral trades that would be Pareto improving and that both parties know about —they know both what the Pareto improving contracts are and who would gain by agreeing to them—but these trades are not carried out. Moreover, (5.2) permits allocations in which agents on the short side of the market receive only a zero payoff (that is, do no better than if they were not trading) whereas agents on the long side receive a positive payoff as well as, in the way one might expect, the other way round. This is true even when $E = \min\{N, L\}$ so that all agents on the short side are participating in a contract. It also has the implication that, in an equilibrium with $E < L$ and thus unemployed workers, all employed workers can receive strictly greater expected lifetime utility than unemployed workers so that, since all workers are identical, the unemployment is genuinely involuntary. This is in the spirit of the involuntary unemployment result in Shapiro and Stiglitz (1984) (which also arises because performance is not verifiable) although, in contrast to Shapiro and Stiglitz, there is not a unique unemployment equilibrium. Nor, in the spirit of Carmichael's (1985) critique of Shapiro and Stiglitz, is involuntary unemployment an essential characteristic of equilibrium. Condition (5.2) permits allocations in which all workers, whether employed or unemployed, receive zero utility so that unemployment is not involuntary. Moreover, it permits allocations in which there are vacancies that are unfilled despite those that are filled earning strictly positive profits—these might be called involuntarily unfilled vacancies—as well as allocations in which both filled and unfilled vacancies generate zero profits.

Even though Proposition 4 is very much in the spirit of the "Folk Theorem" results, one might still ask, first, whether the rules of the game that we specified above, and that govern the operation of the labor market, are really the appropriate ones and, second, whether the equilibrium concept we have used discriminates adequately what one might regard as "reasonable" equilibria from "unreasonable" ones. On the first question, the obvious candidate for a change in the rules of the game is to have workers, rather than firms, announcing the contracts. It is clear, however, that Proposition 4 applies unchanged to this alternative game. The necessity part of the proof is independent of who announces the contracts and an obvious respecification of the strategies in the sufficiency part leaves the conclusion unchanged.

The question of the appropriate equilibrium concept is more tricky. We deal here with one issue, that of unfilled jobs and unemployed workers co-existing despite there being mutually beneficial self-enforcing contracts, but leave to the next section the issue of possible renegotiation once a contract has been terminated. On that first issue, however unreasonable it may seem from the standpoint of traditional economic models, it does not seem to us so unreasonable in the light of the contracting problems that result from performance not being verifiable. In the present model, producing positive output requires a degree of trust by at least one of the parties that the other will stick to a contract. A worker will not incur disutility to produce positive output in a job unless he or she believes that the employer will adhere to the implicit contract by paying any agreed bonus and continuing the contract in the next period. Similarly, a firm will not keep on a worker, nor pay a bonus that is not legally binding, unless it believes that the worker will adhere to the implicit agreement by continuing the contract and producing the agreed output in the next period. If one party believes the other will not adhere to a contract, and the other believes that the first party believes this, then not adhering to that contract is optimal for both and it will not be offered in the first place. This is, in principle, exactly as in other models with multiple Nash equilibria. A particular equilibrium is played only if all players have appropriate beliefs about the strategies other players will play—see Aumann (1987). In this respect, the present market game is like a repeated prisoners' dilemma game—an efficient outcome can be an equilibrium if the parties have the right beliefs but there is no guarantee that they will. The essence of the proof of Proposition 4 is that, in implementing a particular market allocation, nobody believes other agents will abide by any implicit contracts other than those that implement that allocation. Thus, it is not optimal for any other contracts to be offered.

In practice, too, it is not unreasonable to imagine that markets that depend on trust, financial markets in the City of London are obvious examples, would not actually come into being, at least in their present form, if they did not have a history that led agents to have that trust. Nor is it unreasonable to imagine economies being underdeveloped in the literal sense that certain markets do not exist because nobody believes that those potentially on the other side of the market will stick to an implicit contract. Proposition 4 says that this can result from everybody following an optimal strategy and that, if some implicit contracts are believed while others would not be, then those contracts are the ones that will be offered even if they are not efficient.

6. EQUILIBRIA WITH RENEGOTIATION

Punishment strategies like those in the proof of Proposition 4 have become standard in models of infinitely repeated games. Yet some economists have doubted whether the punishments involved are really credible. To punish agents who deviate from equilibrium contracts, other agents forego the potential gains from contracting with them in the future. For these punishments to be credible, those agents must be prepared to deny themselves the possibility of renegotiating deals that would reap these gains. Since bygones are bygones, it is not clear that it is reasonable to assume that they will do so. But if they renegotiate, the punishments implicit in the original contract are not actually carried out and rational agents should take account of that in deciding what contracts will be self-enforcing in the first place. To be self-enforcing a contract would then need to be proof against renegotiation of this sort.

In this section we consider the implications of renegotiation for market equilibrium in the model. Pearce (1987) has suggested one notion of renegotiation proofness and explores its structure formally. Here we concentrate on its implications for the present market game. The essence of the notion is as follows. Suppose a game is in a punishment phase. If players renegotiate away from that punishment, it is not internally consistent to renegotiate to an allocation that requires the original punishment to sustain it. Sustaining an allocation by a punishment entails believing that it is not possible to renegotiate away from that punishment and it is internally inconsistent to believe that, when the opportunities available are stationary over time, a punishment that one is just renegotiating away from will not be renegotiated away from in future. For a symmetric equilibrium, the logical conclusion is that rational players should not renegotiate to any allocation that requires punishments as stringent to sustain it as the punishments they are negotiating away from. If it did, that allocation would not itself be renegotiation proof.

The present market game has an important asymmetry, however, at least as long as the number of jobs is not the same as the number of workers. In any renegotiation, there will then be a short side and a long side of the market. There is nothing incredible in the agents on the long side of the market having punishment payoffs of zero even after renegotiation. Suppose, for example, there are more workers than jobs. A firm that loses a worker will then always be able to find another identical worker as a replacement. Since in equilibrium no jobs are ever actually vacated, it seems reasonable that any worker who deviates from an equilibrium strategy and thus loses a job, will never get another one. Then that worker's remaining lifetime utility will be zero. A symmetric argument applies if there are more jobs than workers.

In the present context, the punishments that follow from cheating on a contract at time t are the default payoffs Π_t^0 and U_t^0. These are the payoffs that a firm and a worker get if their contract comes to an end and they seek alternatives on the market. For agents on the long side of the market, the argument above implies that their default payoffs are zero. To see the implications of Pearce's approach to renegotiation proofness for the short side, consider the case $L \geqslant N + 1$ so that there are more workers than jobs. Suppose a firm is negotiating a new contract in some period $t \geqslant 2$ in a punishment phase of the game. Its default payoff in period t is, by definition, the payoff it gets from this renegotiation. In carrying out the renegotiation, both parties must bear in mind that exactly the same renegotiation could take place between the firm and another worker in the future. Thus, to be proof against such further renegotiation, the contract negotiated at t must be sustained by a credible punishment. That is, the punishment must correspond to a self-enforcing allocation that is itself renegotiation proof. Let x^* denote the allocation that gives the most stringent of all future credible punishments on the firm from period 2 on and let $\Pi_t(x^*)$ be the payoff from period t on if the punishment implied by x^* is started at t. Any allocation worse than x^* for the firm cannot, by the definition of x^*, be self-enforcing and renegotiation proof, so $\Pi_t(x^*)$ is a lower bound on the firm's payoff at t for any self-enforcing and renegotiation proof allocation. Hence, for a renegotiation proof market equilibrium $\Pi_t^0 \geqslant \Pi_t(x^*)$, all $t \geqslant 1$. Certainly, therefore, a necessary condition for x^* to be self-enforcing is that it be self-enforcing for the default payoff $\Pi_t(x^*)$ at each t; that is, $x^* \in X^*[\Pi(x^*), \varnothing]$, where $\Pi(x^*) \equiv [\Pi_1(x^*), \Pi_2(x^*), \ldots]$ is the sequence of payoffs with the tth term the firm's payoff if x^* is started at t and $\varnothing$ is the sequence of default payoffs of zero that workers have when they are on the long side of the market. For x^* also to be renegotiation proof, it must maximize $\Pi_t(x^*)$ for all $x \in X^*[\Pi(x), \varnothing]$. Otherwise there exists a self-enforcing allocation better for the firm that requires a punishment less stringent than x^*. But then there cannot be an allocation better for the firm than x^* that is self-enforcing, renegotiation proof, and also supported by a punishment less severe than x^*. Moreover, any allocation that has a punishment more severe than x^* is not, but the definition of x^*, self-enforcing and renegotiation proof. Hence, for any self-enforcing contracts to be renegotiation proof, the default payoffs must be given by $\Pi_t^0 = \Pi_t(x^*)$, all $t \geqslant 1$. Again, a symmetric argument applies when there are more jobs than workers. This motivates the following definition.

DEFINITION: A market allocation $(x^1, \ldots, x^E) \in X^E$ can be implemented by a *renegotiation proof equilibrium* if

$$(6.1) \qquad x^i \in X^*(\Pi^0, U^0) \qquad (i = 1, \ldots, E),$$

$$(6.2) \qquad \Pi_t(x^i) \geqslant \Pi_t^0 \ (i = 1, \ldots, N), \text{ and } U_t(x^i) \geqslant U_t^0 \quad (i = 1, \ldots, L; \text{ all } t),$$

when Π_t^0 and U_t^0 satisfy:

(a) for $L \geqslant N + 1$: $U_t^0 = 0$ and $\Pi_t^0 = \Pi_t(x^*)$, all t, where x^* satisfies (i) $x^* \in X^*[\Pi(x^*), \varnothing]$, (ii) there does not exist $x' \in X^*[\Pi(x'), \varnothing]$ such that $\Pi_t(x') > \Pi_t(x^*)$ for any $t \geqslant 2$;

(b) for $N \geqslant L + 1$: $\Pi_t^0 = 0$, and $U_t^0 = U_t(x^*)$, all t, where x^* satisfies (i) $x^* \in X^*[\varnothing, U(x^*)]$, (ii) there does not exist $x' \in X^*[\varnothing, U(x')]$ such that $U_t(x') > U_t(x^*)$ for any $t \geqslant 2$.

In this definition, (6.1) ensures that the allocation can be implemented by self-enforcing contracts and (6.2) that all agents receive the appropriate default payoff. Conditions (a) and (b) ensure that, for the cases $L \geqslant N + 1$ and $N \geqslant L + 1$ respectively, the default payoffs can be sustained by the allocation x^* that is itself self-enforcing and renegotiation proof. We do not treat the case $N = L$ since there is not then a short side of the market in any future renegotiation and so one must address explicitly a bargaining problem that we do not wish to get into here.

We show in Proposition 5 below that the default payoffs Π_t^0 and U_t^0 must be stationary for an equilibrium to be renegotiation proof. With stationary defaults, Assumption 3 holds so the results in Proposition 2 apply. To ensure that (4.18) can be satisfied, and so employment occur, with the firm or the worker having a positive default payoff, we make the following assumption.

ASSUMPTION 5: $\mathrm{Max}_{p_t \geqslant 0}[y(p_t) - v(p_t)/\delta] > 0$.

It is convenient to define

$$p^* = \arg\max_{p_t \geqslant 0}\left[y(p_t) - v(p_t)/\delta\right]. \tag{6.3}$$

PROPOSITION 5: *Under Assumptions* 1, 2, 4, *and* 5, *a market allocation* $(x^1,\dots,x^E) \in X^E$ *can be implemented by a renegotiation proof equilibrium if and only if* $E = \min\{N, L\}$ *and* $x^i \in X^*(\Pi^0, U^0)$, $i = 1,\dots,E$, *when*

(a) *for* $L \geqslant N+1$: $U_t^0 = 0$, *all t, and*

$$\Pi_t^0 = \left[y(p^*) - v(p^*)/\delta\right]/(1-\delta), \quad \textit{all } t; \tag{6.4}$$

(b) *for* $N \geqslant L+1$: $\Pi_t^0 = 0$, *all t, and*

$$U_t^0 = \left[y(p^*) - v(p^*)/\delta\right]/(1-\delta), \quad \textit{all } t. \tag{6.5}$$

PROOF: (a) *Necessity*. Consider first the case $L \geqslant N+1$. Let $(x^1,\dots,x^E)$ be a market allocation implemented by a renegotiation proof equilibrium under the assumptions stated. That $x^i \in X^*(\Pi^0, U^0)$ and that $U_t^0 = 0$, all t, follow directly from the definition of a renegotiation proof equilibrium. It also follows from (ii) of that definition that x^* must be such that

$$\Pi_t^0 = \max_x\left\{\Pi_t(x) | x \in X^*[\Pi(x), \varnothing]\right\}, \quad \text{all } t. \tag{6.6}$$

Since this is constant through time, we can define $\pi^0 = (1-\delta)\Pi_t^0$ and $u^0 = (1-\delta)U_t^0$, all t. Thus Assumption 3 is satisfied and Proposition 2 applies. But, by Assumption 5, (4.18) is satisfied for some $\pi^0 > 0$ so, by (a) of Proposition 2, the program in (6.6) is feasible. It then follows from (4.18) and $u^0 = 0$ that its solution satisfies $\pi^0 = [y(p^*) - v(p^*)/\delta] > 0$, with p^* defined in (6.3). Thus, since $\Pi_t^0 = \pi^0/(1-\delta)$, (6.4) holds. Finally, since all jobs must have a payoff of at least Π_t^0, it must be that $E = N$ since otherwise some job receives payoff sequence $\varnothing$. The proof for the case $N \geqslant L+1$ is precisely symmetric with the program

$$U_t^0 = \max_x\left\{U_t(x) | x \in X^*[\varnothing, U(x)]\right\}, \quad \text{all } t, \tag{6.7}$$

replacing that in (6.6).

(b) *Sufficiency*. Again, consider the case $L \geqslant N+1$ first. Clearly, the x^* solving (6.6) satisfies (i) of the definition of a renegotiation proof equilibrium and, in view of (6.7), $\pi^0 = (1-\delta)\Pi_t^0$, with Π_t^0 defined by (6.4), is the highest feasible stationary value of profit. Thus this x^* also satisfies (ii) of the definition of a renegotiation proof equilibrium. That $x^i \in X^*(\Pi^0, U^0)$ implies trivially that (6.1) is then satisfied. That $E = N$ ensures that $\Pi_t(x^i) \geqslant \Pi_t^0$, $i = 1, \ldots, N$ and all t. That $U_t^0 = 0$ ensures $U_t(x^i) \geqslant U_t^0$, $i = 1, \ldots, L$ and all t. Hence (6.2) is also satisfied. The proof for $N \geqslant L + 1$ is symmetric. *Q.E.D.*

This proposition characterizes the set of allocations that can be implemented by renegotiation proof equilibria. In contrast to the result in Proposition 4, either all jobs are filled or all workers employed. The reason is that an allocation that leaves an agent on the short side of the market without a partner (and, therefore, receiving a payoff of zero) cannot be renegotiation proof because there is always an unmatched agent on the long side of the market with whom to renegotiate a positive payoff. The market then works to raise the payoff of agents on the short side in any renegotiation to the highest level consistent with renegotiated contracts being renegotiation proof. That is what (6.4) and (6.5) specify. That these are the highest payoffs possible follows directly from (a) of Proposition 2 and the definition of p^* in (6.3)—a higher value of Π_t^0 or U_t^0 would violate (4.18). Note that $p^* < p^0$, the efficient level of performance defined in (2.5) that would be set if performance were verifiable, so the nonverifiability of performance has a real cost even though either all jobs are filled or all workers employed. The market cannot get to the Pareto frontier that could be achieved if performance were verifiable.

Proposition 5 specifies unique sequences of default payoffs for any renegotiation proof equilibrium but there is not a unique equilibrium allocation. The reason for this should be clear from the earlier discussion of self-enforcing contracts. Condition (4.4) of Proposition 1 required that, for a contract to be self-enforcing, it must offer a strictly positive surplus over the default payoffs from period 2 on. But, with a stationary model and stationary default payoffs, it must then be possible to generate a strictly positive surplus in period 1 also. With the stationary level of performance p^*, it follows from the definition in (2.5) that this surplus is given by

$$(6.8) \qquad [y(p^*) - v(p^*)]/(1-\delta) - U_1^0 - \Pi_1^0 = v(p^*)/\delta > 0,$$

the equality following from substitution of the default payoffs in either (a) or (b) of Proposition 5. This surplus is like a rent. It makes no difference to incentive compatibility whether it goes to the firm or to the worker. It can, in fact, go to either or even be dissipated by the adoption of a contract with performance less than p^*.

One way to generate a unique equilibrium is to follow the approach to selecting equilibria adopted in Farrell (1985). To see its implications in the present context, recall that a contract $\omega^i \equiv (s^i, \sigma^i, c^i)$ offered for job i specifies not only the explicit, legally enforceable part c^i but also suggestions s^i and σ^i as to the strategies that a worker and a firm adopting this contract should play. There is no obligation for them actually to play these strategies—there cannot be because they are not legally enforceable. Farrell, however, makes the assumption that players believe the suggestions *will* actually be played provided they are consistent in an appropriate sense. In the present context, a natural interpretation of consistency is that the allocation implemented by the suggested strategies is self-enforcing and renegotiation proof. If both workers and firms believe that their partner will actually play the suggested strategy, the best response of workers in period 1 is to accept the contract that, of all the self-enforcing and renegotiation proof contracts on offer, yields the highest utility if adhered to. To express this formally, recall that Υ_l^1 was defined as the set of jobs still unfilled when worker l makes the choice of what contract to accept, or chooses to remain unemployed, in period 1. Now define Υ_l' as the set of jobs in Υ_l^1 for which the contracts offered are self-enforcing and renegotiation proof. Formally,

$$\Upsilon_l' \equiv \{ i \mid i \in \Upsilon_l^1, \omega^i \equiv (s^i, \sigma^i, c^i) \in \Omega^*(\Pi^0, U^0) \text{ and } x^i \in X^*(\Pi^0, U^0) \text{ when } \Pi^0 \text{ and } U^0 \text{ satisfy the definition of a renegotiation proof equilibrium} \}.$$

ASSUMPTION 6: *Firms believe that, in period* 1, *worker* l, $l = 1, \ldots, L$, *will accept the contract for job i that solves* $\max_{i \in \Upsilon_l'} U(x^i)$ *and play strategy* s^i. *Worker* l *believes that, if in period* 1 *he or she accepts contract* ω^i *for job* $i \in \Upsilon_l'$, *then the firm with job i will play* σ^i.

COROLLARY 1: *Under Assumptions* 1, 2, 4, 5, *and* 6, *a market allocation* $(x^1, \ldots, x^E) \in X^E$ *can be implemented by a renegotiation proof equilibrium if and only if* $E = \min\{N, L\}$, $x^i \in X^*(\Pi^0, U^0)$ *for* Π^0 *and* U^0 *as specified in Proposition* 5, *and*

(a) *for* $L \geqslant N + 1$: $U_1(x^i) = 0$, $i = 1, \ldots, N$, *and*

$$\Pi_1(x^i) = [y(p^*) - v(p^*)]/(1 - \delta) \qquad (i = 1, \ldots, N); \tag{6.9}$$

(b) *for* $N \geqslant L + 1$: $\Pi_1(x^i) = 0$, $i = 1, \ldots, L$, *and*

$$U_1(x^i) = [y(p^*) - v(p^*)]/(1 - \delta) \qquad (i = 1, \ldots, L). \tag{6.10}$$

PROOF: The proof for the case $L \geqslant N + 1$ follows directly from the fact that, under Assumption 6, the firm with job i will offer a self-enforcing contract that implements an x^i that satisfies

$$x^i \in \arg\max_x \{ \Pi_1(x) \mid x \in X^*(\Pi^0, \varnothing) \text{ when } \Pi^0 \text{ satisfies (6.4), all } t \}.$$

But then the firm receives all the output except what it has to pay the worker to just compensate for the disutility of effort, giving the profit in (6.9). For $N \geqslant L + 1$, a symmetric argument establishes that the worker gets all the output, giving the lifetime utility in (6.10). *Q.E.D.*

This outcome is Walrasian in the sense that agents on the short side of the market get all the potential surplus from trading, whereas agents on the long side get only the reservation payoff of zero that they could achieve if they withdrew from the market. Note that, from Proposition 3, these payoffs can be achieved by a stationary allocation (and hence a constant positive wage in each period) without the need for severance pay or bonding. Note also, however, that

Assumption 6 is an *assumption*. We do not see that it follows necessarily from the logic of rational strategic choice. Like the notion of renegotiation proofness, it restricts the set of beliefs about the strategies other players will play and we know from Aumann (1987) that any correlated equilibrium, and so certainly any Nash equilibrium, is consistent with Bayesian rationality for some set of beliefs about the strategies other players will adopt. Indeed, the kind of restriction imposed by Assumption 6 in the case $L \geqslant N + 1$ would ensure that, in any two-player game with multiple equilibria, if one player suggests playing a particular equilibrium, then that equilibrium will be played.

Theorists in the efficiency wage tradition have argued that, in equilibrium in models of this type, employed workers get strictly higher utility than unemployed workers.[4] As Corollary 1 makes clear, however, this is not a necessary characteristic of equilibria in this model and it is, therefore, incorrect to say that the limitation on enforceable contracts that results from performance being unverifiable by itself necessarily implies such an outcome. But there are beliefs different from those in Assumption 6 that can result in an efficiency wage outcome. Suppose agents on the long side of the market believe that, because they are in a stationary environment, those on the short side believe they will always be able to negotiate as good a contract in the future as they can in the present. Then no punishment worse than the present contract is credible and a contract will be self-enforcing only if it requires a punishment no more severe than the allocation it implements. In a sense, this just takes one step further the reasoning behind the definition of renegotiation proofness. Renegotiation proofness implies that, once one contract has been terminated, other agents will agree to another contract with either of the parties involved only if that contract requires a punishment no more severe than the allocation it implements. The beliefs just described imply that the same condition applies to contracts in period 1, before any cheating has taken place. The initial allocations that are regarded as credible in period 1 are then just those that would be regarded as credible if negotiated at a later date. These beliefs are expressed formally as follows.

ASSUMPTION 7: *Suppose in period* 1 *a contract with allocation* x^i *is offered and accepted for job* i, $i = 1, \ldots, E$. *If* $L \geqslant N + 1$, *workers believe that the firm with job* i *believes* $\Pi_t^0 = \Pi_1(x^i)$, *all* $t \geqslant 2$. *For* $N \geqslant L + 1$, *firms believe that the worker accepting job* i *believes that* $U_t^0 = U_1(x^i)$, *all* $t \geqslant 2$. *If two or more contracts give the same profit, the firm offers that giving the highest utility to the worker.*

COROLLARY 2: *Under Assumptions* 1, 2, 4, 5, *and* 7, *a market allocation* $(x^1, \ldots, x^E) \in X^E$ *can be implemented by a renegotiation proof equilibrium if and*

[4] Many efficiency wage models follow Malcomson (1981) and Shapiro and Stiglitz (1984) in assuming that firms can monitor a worker's effort only imperfectly. The role of imperfect monitoring in many of those models, however, is merely to ensure that, when time is treated continuously, there is a strictly positive probability of a delay between shirking taking place and the worker getting dismissed as a result. Thus there is a positive probability of a worker receiving a wage for some time after deciding to shirk. Treating hiring periods as discrete, as we do here, has the same effect.

only if $E = \min\{N, L\}$, $x^i \in X^*(\Pi^0, U^0)$ *for* Π^0 *and* U^0 *as specified in Proposition* 5, *and* (a) *for* $L \geqslant N+1$:

$$\Pi_1(x^i) = [y(p^*) - v(p^*)/\delta]/(1-\delta) \qquad (i=1,\ldots,N), \tag{6.11}$$

$$U_1(x^i) = v(p^*)/\delta \qquad (i=1,\ldots,N);^5 \tag{6.12}$$

(b) *for* $N \geqslant L+1$:

$$U_1(x^i) = [y(p^*) - v(p^*)/\delta]/(1-\delta) \qquad (i=1,\ldots,L), \tag{6.13}$$

$$\Pi_1(x^i) = v(p^*)/\delta \qquad (i=1,\ldots,L). \tag{6.14}$$

Conditions (6.11) and (6.13) follow directly from applying Assumption 7 to (6.4) and (6.5) of Proposition 5 respectively. They specify for the two cases the payoffs to agents on the short side of the market, which are limited to the default payoffs by the beliefs in Assumption 7. Thus, they receive no surplus over and above their default payoff. Conditions (6.12) and (6.14) specify the payoffs to agents on the long side. For $N \geqslant L+1$ so that there are more jobs than workers, (6.14) specifies that jobs with workers get the payoff $v(p^*)/\delta$, which is the maximum profit possible given the payoffs to workers in (6.13). This profit is just the surplus given in (6.8). Since it is strictly positive and since jobs without workers have zero profits, unfilled vacancies (of which, with more jobs than workers, there must be some) are then necessarily involuntarily unfilled. For $L \geqslant N+1$ so that there are more workers than jobs, (6.12) specifies that workers with jobs get lifetime utility $v(p^*)/\delta$, which is the maximum possible given the payoff to firms in (6.11). Again, this is just the surplus given in (6.8). Clearly, with more workers than jobs, there has to be unemployment. But the lifetime utility of workers with jobs is strictly positive, whereas that of workers without jobs is zero, so the unemployment is involuntary.

[5] The part of Assumption 7 specifying that, if two or more contacts give the same profit, the firm offers that giving the highest utility to the worker, is important for the uniqueness of the lifetime utility in (6.12) but not for any other aspect of Corollary 2. Without it, the lifetime utility in (6.12) could be anywhere in the interval $[0, v(p^*)/\delta]$ and need not be the same in every job. The reason is that, as we have specified the rules of the market game, it is the firm that offers the contract and it can offer one that reduces the surplus without affecting its profit by offering an allocation with $p_1 < p^*$ and paying a wage that is lower by precisely the same amount as revenue is reduced. This reduces utility but still leaves the contract self-enforcing because that requires the appropriate surplus only from period 2 on. With $L \geqslant N+1$, some worker will accept such a contract as long as it gives at least as much utility as unemployment, so the firm need offer no more than zero. A dominance argument is not sufficient to ensure that the firm offers a contract yielding $U_1(x^i) = v(p^*)/\delta$ because that requires it to commit itself to paying a higher wage if the worker were to adopt the nonequilibrium strategy of setting $p_1 = 0$. Thus any utility level between 0 and $v(p^*)/\delta$ can be an equilibrium. This nonuniqueness is not, however, robust to changes in the specification of the rules of the market game. If we had specified that workers offered the contracts instead of firms, the resulting market equilibrium would still have the level of profit specified in (6.11) but the workers would choose those contracts that maximized their utility for this level of profit. Then only the utility level $U_1(x^i) = v(p^*)/\delta$, all i, would be an equilibrium. We do not, therefore, attach much significance to the other possible equilibria.

7. CONCLUSION

The equilibrium in Corollary 2 when there are more workers than jobs is essentially the efficiency wage unemployment equilibrium of Shapiro and Stiglitz (1984)—performance is stationary through time and employed workers receive a surplus from employment. This corollary, therefore, provides a theoretical foundation for such an equilibrium. But modelling contracts in the general way we have done here, rather than assuming that they take a particular form as Shapiro and Stiglitz do, makes it clear that it is not the need for contracts to be self-enforcing that prevents unemployed workers undercutting employed workers by offering to work under a contract that gives firms higher profits. That follows from Propositions 4 and 5 and Corollary 1. Equally, extending contracts to allow bonding, as suggested by Carmichael (1985), does not ensure that all unemployment is voluntary in equilibrium. Whether or not bonding is permitted makes no difference to the set of equilibria in the model, as Proposition 3 has shown. What is crucial in determining whether or not involuntary unemployment occurs is the nature of beliefs. With more jobs than workers, a characteristic feature of the equilibria of Proposition 5 is that, for any contract giving profit greater than that in Corollary 2, the default payoff required to sustain that contract is lower than the profit it generates. If workers believe that, by accepting a contract with such high profit in period 1, the firm will be led to believe that it can get a default payoff of this amount in the future, then it is not rational for them to do so. That is what Corollary 2 shows. Under other beliefs, as Corollary 1 has shown, unemployment is not involuntary.

The unemployment of Corollary 2 is involuntary in the ex ante sense that unemployed workers have expected lifetime utility strictly lower than those currently being hired. This is different from the ex post sense familiar in contract theory, in which all workers get the same expected lifetime utility in the initial labor market but, because contracts offer only partial insurance against productivity shocks, those unlucky enough to be laid off have remaining lifetime utility lower than those retained. Thus it provides a theory of involuntary unemployment rather than one of involuntary layoffs. But note that the equilibria are symmetric. When there are more jobs than workers, firms have vacancies that are involuntarily unfilled.

Assumptions 6 and 7 are not the only restrictions that one might impose on beliefs to select from the equilibria of Proposition 5. We have studied these because they focus attention on the kinds of restrictions that result in, respectively, an outcome that is Walrasian in the sense that agents on the short side of the market get all the surplus and an efficiency wage outcome of the type analyzed in Shapiro and Stiglitz (1984). We would, however, emphasize that the results of Proposition 4 reveal difficulties for the functioning of markets in which performance is not verifiable beyond those raised by efficiency wage models. In both the efficiency wage and the Walrasian outcomes derived above, either all workers are employed or all jobs filled. But, as Proposition 4 makes clear, that is not a necessary consequence of the model. It depends on agents believing that other agents will actually carry out some self-enforcing implicit contract. Without that, trades that are potentially mutually beneficial may never take place. This belief is implicit in the concept of renegotiation proofness we have used. But we regard that concept as a statement of what might be reasonable under certain circumstances, not as something inherent in the notion of rational strategic choice.

The central message from this is that beliefs are crucial in determining the equilibria in such markets. One cannot determine the outcome simply from the technology and the utility functions of agents. Since the problem of verifying performance seems to be widespread in labor markets (as indeed in other markets), this conclusion has practical, as well as theoretical, significance. To achieve a better understanding of these types of markets, it is crucial to investigate the nature of the beliefs that are held by agents operating in them.

Department of Economics, Queen's University, Kingston, Ontario, Canada K7L 3N6

and

Department of Economics, University of Southampton, Southampton SO9 5NH, U.K.

Manuscript received January, 1985; final revision received April, 1988.

REFERENCES

ALCHIAN, ARMEN A., AND HAROLD DEMSETZ (1972): "Production, Information Costs, and Economic Organization," *American Economic Review*, 62, 777–795.

AUMANN, ROBERT J. (1974): "Subjectivity and Correlation in Randomized Strategies," *Journal of Mathematical Economics*, 1, 67–96.

——— (1987): "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica*, 55, 1–18.

AZARIADIS, COSTAS (1975): "Implicit Contracts and Underemployment Equilibria," *Journal of Political Economy*, 83, 1183–1202.

BULL, CLIVE (1987): "The Existence of Self-Enforcing Implicit Contracts," *Quarterly Journal of Economics*, 102, 147–159.

CARMICHAEL, LORNE H. (1983): "Firm-Specific Human Capital and Promotion Ladders," *Bell Journal of Economics*, 14, 251–258.

——— (1985): "Can Unemployment Be Involuntary? Comment," *American Economic Review*, 75, 1213–1214.

EATON, B. CURTIS, AND WILLIAM D. WHITE (1982): "Agent Compensation and the Limits of Bonding," *Economic Inquiry*, 20, 330–343.

——— (1983): "The Economy of High Wages: An Agency Problem," *Economica*, 50, 175–181.

FARRELL, JOSEPH (1985): "Communication and Nash Equilibrium," GTE Labs Working Paper 85-10.

FUDENBERG, DREW, AND ERIC MASKIN (1986): "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, 54, 533–554.

GROSSMAN, SANFORD J., AND OLIVER D. HART (1981): "Implicit Contracts, Moral Hazard, and Unemployment," *American Economic Review*, 71, 301–307.

HART, OLIVER D., AND BENGT HOLMSTROM (1987): "The Theory of Contracts," in *Advances in Economic Theory: Fifth World Congress*, ed. by Truman Bewley. Cambridge: Cambridge University Press, 71–155.

HOLMSTROM, BENGT (1981): "Contractual Models of the Labor Market," *American Economic Review*, 71, 308–313.

KOHLBERG, ELON, AND JEAN-FRANCOIS MERTENS (1986): "On the Strategic Stability of Equilibria," *Econometrica*, 54, 1003–1037.

MACLEOD, W. BENTLEY, AND JAMES M. MALCOMSON (1988): "Reputation and Hierarchy in Dynamic Models of Employment," *Journal of Political Economy*, 96, 832–854.

——— (1989): "Labour Turnover and the Natural Rate of Unemployment: Efficiency Wage vs. Frictional Unemployment," Centre for Labour Economics Discussion Paper No. 336, London School of Economics.

MALCOMSON, JAMES M. (1981): "Unemployment and the Efficiency Wage Hypothesis," *Economic Journal*, 91, 848–866.

——— (1984): "Work Incentives, Hierarchy, and Internal Labor Markets," *Journal of Political Economy*, 92, 486–507.

——— (1986): "Rank-Order Contracts for a Principal with Many Agents," *Review of Economic Studies,* 53, 807–817.

PEARCE, DAVID G. (1987): "Renegotiation-Proof Equilibria: Collective Rationality and Intertemporal Cooperation," mimeo, Yale University.

RADNER, ROY (1985): "Repeated Principal-Agent Games with Discounting," *Econometrica,* 53, 1173–1198.

SELTEN, R. (1975): "Re-examination of the Perfectness Concept for Equilibrium Points in Extensive Games," *International Journal of Game Theory*, 4, 25–55.

SHAPIRO, CARL, AND JOSEPH E. STIGLITZ (1984): "Equilibrium Unemployment as a Worker Discipline Device," *American Economic Review*, 74, 433–444.

SIMON, HERBERT A. (1957): *Models of Man*. New York: Wiley.

THOMAS, JONATHAN, AND TIM WORRALL (1988): "Self-Enforcing Wage Contracts," *Review of Economic Studies*, 55, 541–554.

VAN DAMME, E. E. C. (1983): "Refinements of the Nash Equilibrium Concept," Ph.D. thesis, Technische Hogeschool Eindhoven.

WILLIAMSON, OLIVER E., MICHAEL L. WACHTER, AND JEFFREY HARRIS (1975): "Understanding the Employment Relation: The Analysis of Idiosyncratic Exchange," *Bell Journal of Economics*, 6, 250–278.