COOPERATION, CORPORATE CULTURE AND INCENTIVE INTENSITY

by

R. ROB*
and
P. ZEMSKY**

97/51/EPS/SM

* Professor of Economics, University of Pennsylvania, USA.

** Assistant Professor of Strategy at INSEAD, Boulevard de Constance, Fontainebleau 77305 Cedex, France.

Cooperation, Corporate Culture and Incentive Intensity

Rafael Rob and Peter Zemsky*

May 14, 1997

Abstract

We develop a theory of the firm in which the willingness of workers to cooperate with each other plays a central role. We study a dynamic principal-agent problem. In each period, the firm (the principal) chooses an incentive intensity (how much to pay workers per-unit of measured output) and the employees (the agents) allocate effort between individual production and tasks that involve cooperating with other employees. Following the literature on organizational behavior, (i) employees are willing to engage in cooperative tasks even when these tasks are less effective at increasing their measured output and (ii) the level of cooperation is increasing in past levels of cooperation in the firm and decreasing in the incentive intensity. Hence, an increase in the incentive intensity does not just increase current effort, it has important dynamic consequences: future employee cooperativeness falls. We show how the firm balances these two effects to maximize its lifetime profits. By extending the set of employee motivators beyond the purely financial, we are able to introduce a precise definition of corporate culture and to show how firms optimally manage their culture. Our theory helps explain why different firms, placed in similar "physical" circumstances, choose different incentive systems. It also helps explain how corporate culture can be a hard-to-imitate asset which yields some firms excess profits.

Keywords: altruism, endogenous preferences, prisoner's dilemma, barriers to imitation, history dependence

*University of Pennsylvania and INSEAD, respectively. We thank numerous seminar audiences for comments on the paper.

1. Introduction

What motivates "organizational" man? For economists the answer is individual self-interest and the system of incentives that comprise the internal organization of the firm. In contrast, researchers in organizational behavior take a much broader view of motivation, which includes social norms and pressures to conform to these norms. For these scholars, organizations are not primarily systems of incentives, but social processes that shape the preferences and behavioral norms of employees (Miller, 1992).

These two different approaches to individual behavior have lead to two disjoint literatures on organizations. The standard reference for the economics of organizations (Milgrom and Roberts, 1992) adopts "the view that many institutions and business practices are designed as if people were entirely motivated by narrow, selfish concerns". In marked contrast, non-economists can be very critical of the use of incentives. There is much less emphasis on self-interest in an organizational context, especially when cooperation and team work are important determinants of the organization's performance. A well-known example in the recent management literature is the article "Why Incentive Plans Can Not Work" (*Harvard Business Review,* 1993) by the management consultant Alfie Kohn. Buttressed by experimental studies in social psychology and field studies of organizations, Kohn makes claims such as: "Rewards buy temporary compliance, so it looks like the problems are solved. It's harder to spot the harm they cause in the long term." And "The surest way to destroy cooperation and, therefore, organizational excellence, is to force people to compete for rewards...But the same result can occur with *any* use of rewards."

Our aim here is to integrate the two approaches to individual behavior. We consider a principal-agent model in which agents are motivated by a combination of self-interest and social concerns. In so doing, we are able to formalize the arguments of those, such as Kohn, who are hostile to the economic theory of organizations (Why are rewards fundamentally detrimental to cooperation in organizations? Why might the use of rewards have long-term negative consequences?) More fundamentally, we are able to incorporate social concerns and corporate culture into the formal theory of the firm and treat them on par with the traditional objectives of firms and workers. This allows us to investigate the links between social concerns and optimal incentives.

An important element of organizational design is the incentive intensity, the extent to which the rewards of employees are linked to their measured performance. In our model, the incentive intensity controls two choices: The total level of effort an individual exerts, and how much of that effort he devotes to helping his co-workers. These two choices are negatively correlated: As a firm increases the

incentive intensity, employees exert more effort, but they are less cooperative with each other. Such reductions in cooperation have long-term consequences because they erode the goodwill of workers towards each other and undermine norms of cooperation. That is, high-powered incentives can undermine the cooperativeness of a firm's corporate culture and reduce its future productivity.

Consequently, the firm faces an intertemporal trade-off in setting the incentive intensity: It can increase its payoff in the present period, but only at the expense of milking its corporate culture and thereby reducing its payoffs in future periods. The net result is that the firm solves a capital-theoretic problem in which it cares about a stream of payoffs and the elements in this stream are intertemporally connected through the stock of corporate culture.[1]

Using this approach, our results are as follows. First we are able to explain "excess volatility" in incentive systems, relative to what standard principal-agent theory predicts. Taken literally, the received theory predicts that incentive intensity depends on fundamentals such as the firm's production technology, the noisiness of the performance measures, and the effort aversion and risk tolerance of employees. When these fundamentals are held constant, there should be no variation in firms' incentive systems. Yet, cross-sectionally, firms in the same industry with similar workforces can have very different incentive intensities. Likewise, intertemporaly, firms often change their incentive systems without apparent change in the workforce or the technology. Both dimensions of volatility represent important puzzles, which our theory helps resolve.

Another important puzzle that our theory is able to resolve is the existence of "barriers to imitation." It is well-known that firms in the same industry, using the same technology and a similar workforce exhibit vastly different rates of profit. What prevents the less profitable firms from discovering and imitating the practices of their more profitable counterparts? Resource-based theories of strategy assert that an effective corporate culture is a hard-to-imitate asset which can lead to superior performance (Barney, 1986). However, this literature does not give a precise definition of "corporate culture," nor does it explain why imitation can be prohibitively costly. Our theory is more explicit: We develop a model of incentives and cooperation and, within it, provide a precise definition of culture. The model shows that some firms may end up with uncooperative cultures and hence low profitability, while others end up with highly cooperative cultures and high profitability. Yet, the low-profitability firms find it too costly to imitate the culture of their high-profitability counterparts.

[1] In a similar vein, Athey et al. (1994) shows that if the productivity of an employee depends on the extent to which his "type" matches those of his co-workers, then the firm's hiring policy solves a dynamic optimization problem as the firm seeks to manage how the diversity of its stock of employees evolves over time.

The paper proceeds as follows. Section 2 spells out some experimental evidence upon which our formulation is based. Section 3 presents and discusses our model. Section 4 looks at how the incentive intensity affects the level of cooperation and firm profits in a single period. Section 5 explains how the firm has a corporate culture and how that culture creates a dynamic optimization problem for the firm. Section 6 solves the firm's optimization problem. (This section can be skipped or skimmed by readers less interested in technical details.) Section 7 states the main results of the paper. Section 8 discusses extending the analysis to other forms of behavior besides cooperation. Section 9 concludes with thoughts for future work.

2. Related Literature and Details of Our Approach

The formal literature on the economics of organizations views incentive intensity as the solution to a moral-hazard, principal-agent problem. The optimal incentive intensity is determined by a trade-off between the need to generate effort and the cost of placing risk on the employee. In an important contribution to this literature, Holmstrom and Milgrom (1991) point out that employees often have multiple tasks and that there may be no performance measures for some of the tasks.[2]

We follow Holmstrom and Milgrom (1991) and assume that employees engage in multiple tasks and that performance measures are incomplete. However, we deviate from the received literature in two ways. The first deviation is that we assume that some of the output from an employee's unmeasured tasks show up in performance measures of *other* employees, which is necessarily the case if all output is attributed to some employee. A common example of tasks which fit this description is time spent helping other employees, when such helping is not measured (or is not fully credited). Another common example is maintaining shared machines or helping to recruit good colleagues. We refer to the act of devoting effort to an unmeasured task as cooperation. The second (and more significant) deviation from the received literature is that employees are not entirely motivated by self-interest. Instead, we assume that employees are partly altruistic in that they have a taste for cooperation. (Actually, our employees experience "guilt" if they do not cooperate).

Economic researchers have tended not to study altruistic preference, especially in applied work on organizations (see, however, Rotemberg, 1994). However, re-

[2]Holmstrom and Milgrom go on to argue that if a firm wants an employee to devote (unoberved) effort to tasks which have no performance measures, then the firm can not reward any performance measures. In their theory, if a firm does set a positive incentive intensity, the level is still determined by the standard trade-off between motivating effort and risk-sharing. We develop a theory where positive levels of incentive intensity are not determined by risk-sharing.

cent experimental research well documents the existence of altruism. For example, Forsythe et al. (1994) studied a dictator game in which one subject divides $5 between himself and another subject, whom he does not meet.[3] For the 45 subjects in the study, an average of $1.10 was given away, with 64 percent of the subjects giving away at least $1. In a striking result, Hoffman et al. (1994) show that the behavior of subjects is more altruistic when the experimenters observed the play of specific subjects than when they did not. The authors conclude that social considerations are pervasive in game playing and that the standard assumption of purely self-regarding preferences is often violated. The standard assumption seems especially questionable in organizational contexts, where acting in purely self-interested ways is often viewed as inappropriate.

We conclude that the standard principal-agent framework—with its assumption of purely self-interested behavior—may only suffice for characterizing the incentives of isolated employees, such as sales people. For employees embedded in organizations, where social concerns loom large and many cooperative tasks fall between the cracks of performance measures, the standard model is incomplete.

Consistent with the experimental evidence in Hoffman et al. (1994), we do not view employees as having static preferences for altruism, but rather tastes which depend on the social context. In particular, we hypothesize that the taste for cooperation increases with exposure to cooperative behavior by other employees: The more cooperative behavior one experiences, the more likely one is to behave cooperatively oneself. Such history dependent preferences have (at least) two sources. The first is generalized reciprocity, where employees seek to reciprocate cooperation they have experienced (even if they do not repay the specific person who cooperated with them). An example of generalized reciprocity is a person who stops to help a stranded motorist because they were once helped in a similar situation. The second source of history-dependent preferences is a desire to conform to social norms, i.e., to behave like others[4]. An example of conforming to social norms is a person who leaves a tip in a restaurant to which he never expects to return. If employees seek to conform to social norms, then experience in the firm shapes preferences for behaving in certain ways because it makes individuals aware of the behavioral norms around them. There is empirical evidence consistent with the effect of social norms on behavior in games. Berg et al. (1995) show that the altruism of subjects in a trust game is affected by knowledge about how past subjects played. A common finding in one shot, n-person prisoner dilemmas

[3]The dictator game is similar to the more commonly studied ultimatum game except that the second subject is passive and has no ability to block a division.

[4]There is also a vast *Social Psychology* literature, including the book *The Social Animal* by Aronson (1972) which documents the tendency of human being to act in a herd-like fashion, i.e., to yield to group pressures even when they don't necessarily agree with group beliefs (or to act differently than they would have if they were to act in isolation.)

is that the more subjects expect others to cooperate, the more likely they are to cooperate themselves (Dawes, 1980).

Our assumption that employees value cooperation and that they respond to social norms leads us into the realm of "corporate culture". In broad terms, the concept relates to the "set of collectively held values, beliefs, and norms of behavior among members of a firm that influences individual employee preferences and behaviors" (Besanko et al., 1996). Beyond such a broad definition, the culture literature lacks a consensus on the nature of the phenomena — at least to the tastes of economists (see Lazear, 1995 for a summary and critique). Culture research seems to be at an early stage of development where it is still wrestling with basic questions. Such questions have been recently articulated by the sociologist O'Reilly (1989): What is culture? When is culture important? What is the process through which cultures are developed and maintained? How can cultures be managed? Our work tries to address these questions within an economic model of the firm.

An alternative source of cooperation in organizations is reputation mechanisms, which can also be related to the phenomenon of corporate culture (Kreps, 1992). Like our theory, reputation models can explain the existence of different levels of cooperation in otherwise similar firms. However, they lack precise predictions about how cultures develop and are maintained. Because of folk-theorem-type results, almost anything can happen. In contrast, our theory makes precise predictions about the dynamic evolution of cooperation and can trace current levels of cooperation back to initial conditions. More importantly, reputation theory says little about how cultures are managed. In contrast, we show how the firm can actively manage its culture via its choice of incentive intensity.

3. The Model

The firm employs a continuum of risk-neutral workers. The size of its workforce is fixed and normalized to 1. Workers are indexed by $i \in [0, 1]$. The firm operates over an infinite number of periods, $t = 1, 2, \ldots$ In order to define a reasonably tractable dynamic model, we specify particular functional forms for many quantities. We believe that our results would hold in a much more general model.

3.1. The Production and Monitoring Technology

Each period, workers make two decisions: How much **total effort** to exert, e, and how to allocate total effort between **individual production**, e_I, and **cooperative production**, e_C. Each worker's decisions must satisfy $e = e_I + e_C$, $e_I, e_C \geq 0$ and $e_C \leq h$. h is the maximum number of hours which a worker can devote to

cooperative production.

The output from individual production is ae_I, whereas the output from cooperative production is e_C. Therefore the marginal productivity of effort devoted to individual production is a, whereas the marginal productivity of effort devoted to cooperative production is either 1 or 0: 1 for $e_C < h$ and 0 for $e_C > h$.[5] We let $\widetilde{e}_C \equiv e_C/h$. Thus $\widetilde{e}_C$ is the fraction (a number between 0 and 1) of h which the worker devotes to cooperative production. The total output of a worker is:

$$Q = ae_I + e_C.$$

While Q is the true output of a worker, the firm is unable to directly observe it: Some of the cooperative production of a worker shows up in the output of his co-workers, while some of the cooperative production of other workers show up in his own output. For example, cooperative production is answering colleagues' questions, helping recruit a new colleague, or writing a software program which is shared by other workers. These activities enhance the productivity of the *entire* workforce, but they cannot be precisely attributed to the workers who engage in them. Consequently, the firm is only able to observe a proxy of each worker's output, call it $\widehat{Q}$, which equals:

$$\widehat{Q} = ae_I + (e_C + hz')/2 + \varepsilon,$$

where z' is the average level of cooperation in the workforce[6] and ε is a mean-zero random error. We refer to $\widehat{Q}$ as the **performance measure**[7], and to z' as the **average cooperation,** the **level of cooperation** or the **social cooperative**

[5]The general idea is that production is subject to diminishing returns, and that diminishing returns set in faster for cooperative production. h is a parameter of the firm's technology, measuring where diminishing returns set in and how important is team production relative to the firm's overall activities.

[6]That is $z' \equiv \int_0^1 \widetilde{e}_C(i)di$ where $\widetilde{e}_C(i)$ is the (fractional) cooperative input of worker i.

[7]A useful way to think about $\widehat{Q}$ is the following. The firm tries to assess each worker's contribution by eliciting the opinion of his supervisor. However, it is notoriously difficult for supervisors to tell whether low output of a subordinate is due to his laziness or to his devoting time to helping others (and similarly, whether high output is due to industriousness or to taking up disproportionate amount of others' time). Consequently, the time spent on collective efforts gets lumped together, resulting in workers getting less than full credit for their cooperative efforts.

In addition to this, there is another, "organizational," reason for the inability of firms to assess the social contributions of their workers: Even if supervisors are able to discern the cooperative efforts of their subordinates, they may be reluctant to *report* them truthfully, for fear of straining relationships with the subordinates. In many instances the supervisor depends on his subordinates to perform tasks which will make the supervisor look good in the eyes of upper management. So to avert revengeful activities by subordinates, supervisors fill out reports that are fairly uniform and fairly complimentary to their subordinates.

norm. According to this formula, each worker receives credit for only half of his cooperative effort but is, at the same time, credited with half of the average cooperative effort of others. (In a prior version of this paper we derived the expressions for Q and $\widehat{Q}$ from an underlying production and matching technology.)

We make the restriction $0.5 < a < 1$. Under this restriction, effort devoted to cooperative production is more productive than that devoted to individual production: From the formula for Q, the marginal productivity of e_I is a which is less than the marginal productivity of $e_C, 1$. So it is optimal for the firm to "instruct" workers to set $e_C = h$. However, effort devoted to individual production has a greater impact on an individual worker's performance measure (the marginal productivity of e_C in $\widehat{Q}$ is only 0.5). Therefore, if workers are paid on the basis of $\widehat{Q}$ and if they are selfish, they set $e_C = 0$. We let $d \equiv a - 0.5$; d is the opportunity cost of shifting effort from individual to cooperative production.

3.2. Employees' Preferences

Employees' utility depends on their expected wages, the disutility of total effort, the taste for cooperation and how much they cooperate. Specifically,

$$U = E[W] - C(e) - g(h - e_C),$$

where E is the expectation operator, W is the wage, $C(e)$ is the cost of total effort, and $g(h - e_C)$ is the disutility ("guilt") from not cooperating.

We assume a quadratic cost of effort:

$$C(e) = \begin{cases} 0 & \text{if } e \leq \overline{e}, \\ c(e - \overline{e})^2/2 & \text{otherwise,} \end{cases}$$

where $\overline{e}$ is a threshold beyond which employees start to experience disutility of effort. To simplify the analysis we assume $\overline{e} > h$, i.e., employees can choose the maximum level of cooperative production without feeling any disutility of effort.[8]

The disutility from not cooperating, $g(h - e_C)$, is proportional to an employee's sense of guilt, parameterized here by g, and to the amount of effort that is not put into cooperative production, $h - e_C$. We assume that $g = sz + \gamma$, where (i) z is the average cooperation of the workforce in the *prior* period, (ii) s is the propensity to respond to social pressures (same for all workers), and (iii) γ is the predisposition to cooperate. γ represents the effect of idiosyncratic forces, so it varies independently over time and across workers.

[8]This assumption allows us to separate an employee's decisions: First the employee chooses total effort, e, then he decides how to allocate it between e_I and e_C; see Lemma 1.

This representation of g, and especially the intertemporal connection between the prior period cooperative norm and this period guilt, reflects the experimental evidence cited in the introduction: People tend to reciprocate both nice and nasty behavior, and to conform to the social cooperative norm, z, which they have experienced.

We assume γ is uniformly distributed on $[0,1]$, and $0 < s < 1$. Therefore, the effect of experience on preferences is not too large relative to the underlying idiosyncracy. Given this parametrization, the sense of guilt, g, is uniformly distributed in the workforce over the interval $[sz, sz+1]$.

3.3. The Firm's Problem

The firm seeks to maximize the discounted sum of expected per-employee profits given its discount rate δ. The expected per-employee profit in one period is $\Pi = pE[Q] - E[W]$ where p is the price of output. We assume that $p > (1+s)/d$, so that the price is high relative to the willingness to cooperate.[9]

The firm maximizes profits by selecting a **compensation system** in each period. Specifically, we assume that wages are a linear function of an employee's performance measure, $W = b + w\widehat{Q}$. We refer to b as the **base wage** and w as the **incentive intensity**. The compensation system must satisfy an individual rationality constraint: In each period, each employee must be assured a level of utility of at least $\overline{u}$ (the value of outside employment). The firm starts with an initial level of cooperation z_0.

The firm is restricted to the use of a single, linear compensation system for the whole workforce. We use linear incentives for three reasons. First, linear incentives are commonly observed, and are easy to understand and administer. Second, they have the advantage of applying uniform incentive pressure over time (Milgrom and Roberts, 1992). Third, they afford analytical convenience: With linear incentives workers choose either the maximum (h) or the minimum (0) cooperative effort (see Lemma 1). This greatly simplifies the firm's maximization problem.[10]

It is not obvious that it is optimal for the firm to have a single incentive

[9]This simplifies the analysis by assuring that with the first best incentive intensity, $w = p$, there is no cooperation; see Lemma 4.

[10]The firm could try to induce a more desirable level of total effort with a nonlinear scheme, for example a forcing contract. However, for *any* contract, workers still face the decision how to divide effort between individual and cooperative production with individual production being more effective at raising their expected performance measure. So the tradeoff which the firm faces here, choosing between currently high output and low future culture or currently low output but high future culture, will be present for *any* compensation scheme. However, as will be seen below, this tradeoff is easier to work with when the incentive scheme is linear.

intensity for its whole workforce. Although the convexity of the cost function favors uniform incentives since that spreads total effort evenly across the workforce, the existence of rents to defectors (see below) seems to favor partitioning the workforce into two groups, one with high-powered incentives and the other with low-powered incentives. However, there are unmodeled costs of having multiple incentive systems. In particular, treating employees differently can create strained relations between groups and it can increase the costs of administering the incentive system.

4. Incentive Intensity and the One Period Payoff

In each period, employees choose their efforts, e_C and e_I, and the firm chooses the compensation system, b and w. Since there is a continuum of employees, a single employee has no effect on the behavior of other employees or the firm and, thus, employees behave as myopic maximizers. On the other hand, the firm can influence the degree of cooperation, z, and is, thus, behaving as a long-run maximizer. To analyze the firm's problem, we characterize first employees' choices, e_I and e_C, as a function of the incentive intensity w and the extent of cooperation in the previous period z. We then express the firm's one period profit as a function of the same two variables.

The first step is to characterize employees' optimal effort. An employee solves the following maximization problem

$$\max_{e_I, e_C} \{b + w[ae_I + (e_C + hz')/2] - C(e) - g(h - e_C)\}. \tag{4.1}$$

The objective is formed by substituting $E[W] = b + wE[\widehat{Q}]$ and $E[\widehat{Q}] = ae_I + (e_C + hz')/2$ into the definition of an employee's utility. Given equation (4.1), the employee's choice of effort can be broken into two independent problems, the choice of total effort level, e, and the choice of how much to cooperate, e_C. (Individual effort is then given by $e_I = e - e_C$.)

Lemma 1 *(i) The optimal total effort satisfies the first order condition $C'(e^*) = aw$, which gives $e^*(w) = aw/c + \overline{e}$. (ii) If $g > wd$, the employee fully cooperates, $e_C = h$, while if $g < wd$, the employee does not cooperate at all, $e_C = 0$.*

Proof: (i) Since $C(e) = 0$ for $e \leq \overline{e}$, the optimal level of total effort, e^*, exceeds $\overline{e}$. Also, the maximum cooperative effort is h which is $< \overline{e}$. Therefore e^* is chosen independently of the choice of e_C. Substituting $e_I = e - e_C$ into the employee's objective function and eliminating constant terms, we obtain:

$$\max_{e_I, e_C} \{awe - C(e) - w[a - 0.5]e_C - g(h - e_C)\}.$$

>From this expression we see that e^* satisfies the first order condition $C'(e^*) = aw$, and substituting $C'(e) = c(e - \overline{e})$ we obtain the expression for $e^*(w)$.

(ii) The objective function of the employee is linear in e_C with a slope equaling $g - wd$. Therefore, the employee chooses either the maximum, h, or the minimum possible level of cooperative effort, 0, depending on whether g is $>$ or $<$ wd. ■

The factors which affect the choice of total effort are the same as in a standard principal-agent problem: Effort increases with the incentive intensity, w, and the productivity of effort, a, and decreases with the convexity of the effort-cost function, c. The choice of cooperative effort depends on how strong is one's sense of guilt relative to the opportunity cost of cooperating: The stronger is the sense of guilt, the more likely is the employee to cooperate. We call workers who choose $e_C = h$, **cooperators** and workers who choose $e_C = 0$, **defectors** (since cooperation is socially beneficial but privately costly, the problem facing workers resembles the prisoner's dilemma.)

Lemma 2 *The level of cooperation in the current period z' is the following function of the incentive intensity and the level of cooperation in the previous period:*

$$z' = f(w, z) = \begin{cases} 1 + sz - wd & if\ wd \in [sz, sz+1] \\ 0 & if\ wd > sz + 1, \\ 1 & if\ wd < sz. \end{cases}$$

Proof: Since the decision to cooperate depends on $g >$ or $< wd$, the proportion of cooperators in the workforce is given by $z' = Prob(g > wd \mid z)$. Since g is uniformly distributed over $[sz, sz+1]$, we obtain the above expression. ■

Lemma 2 shows the factors which affect the level of cooperation, z', in the current period: The higher is z, the more cooperation employees experience in the previous period, which propels them to cooperate in this period. Hence, a higher z increases the current level of cooperation, z'. On the other hand, the higher is w, the incentive intensity, the bigger is the sacrifice from putting effort into cooperative rather than individual production. Hence, the level of cooperation decreases with incentive intensity.

The next step is to use the above characterization of employees' effort to obtain an expression for the firm's one period profit.

Lemma 3 *The firm's one period (per employee) payoff is the following function of the incentive intensity and the level of cooperation in the previous period,*

$$\Pi(w, z) \;=\; I(w) + B(w, z) - R(w, z),$$

where:

$$I(w) \equiv \frac{a^2}{c} w(p - \frac{w}{2}) + a\bar{e}p,$$

$$B(w,z) \equiv p(1-a)hz',$$

$$R(w,z) \equiv \begin{cases} \overline{u} + wdh(1-z') \text{ if } z' > 0 \\ \overline{u} \text{ if } z' = 0 \end{cases}$$

and

$$z' = f(w,z).$$

Proof: Recall that the one period profit is defined to be $\Pi = pE[Q] - E[W]$. The expected output of an employee is

$$E[Q] = ae^* + (1-a)hz'.$$

The first term, ae^*, is the output if all effort is put into individual production. The second term is the increase in output from shifting hz' units of effort into cooperative production. All that remains is to find an expression for $E[W]$.

The expected wage is $E[W] = b + wE[\widehat{Q}]$. We can eliminate the base wage, b, from this expression as follows. The utility for an employee who cooperates is

$$U_C = b + w(ae^* + hz'/2 - hd) - C(e^*),$$

whereas the utility for one who defects is

$$U_D = b + w(ae^* + hz'/2) - hg - C(e^*).$$

An employee decides to defect whenever $U_D > U_C$, i.e., whenever $hg < hwd$. Therefore, the utility for a defector is higher than the utility for a cooperator. Thus, if the firm tries to maximize profits and if $z' > 0$, the individual rationality constraint on cooperators is binding, while the individual rationality constraint on defectors is not, i.e., $U_C = \overline{u}$. This yields $b = \overline{u} + C(e^*) - w(ae^* + hz'/2 - hd)$. The expected wage is then

$$E[W] = b + w[z'\widehat{Q}_C + (1-z')\widehat{Q}_D] = \overline{u} + C(e^*) + wdh(1-z'),$$

where $\widehat{Q}_C$ ($\widehat{Q}_D$) is the performance measure of cooperators (defectors). On the other hand, if $z' = 0$, there is a discontinuity in the expected wage: The firm no longer has to satisfy $U_C \geq \overline{u}$ because there are no cooperators. Now it can set $U_D = \overline{u}$ and the expected wage is then $E[W] = \overline{u} + C(e^*)$. Combining the expressions for $E[Q]$ and $E[W]$ and substituting the expression for e^* from Lemma 1 gives the result. ■

As a result of this representation the following properties hold.

Lemma 4 *(i) The function $I(w)$ is increasing over $[0,p]$ and decreasing for $w > p$. Therefore it is maximized at $w = p$. (ii) The level of cooperation z' is 0 when $w \geq (sz+1)/d$ and, in particular, when $w = p$. (iii) The function $B(w,z)$ is decreasing in w and increasing in z. (iv) The function $R(w,z)$ is increasing in w and decreasing in z.*

Proof: (i) $I(w)$ is quadratic with a negative coefficient on w^2 and with a maximum at p. (ii) According to Lemma 2, z' vanishes whenever $wd > sz+1$. Since $p > (1+s)/d$, $z' = 0$ when $w = p$. (iii) $B(w,z)$ depends on z' which decreases in w and increases in z.(iv) $R(w,z)$ depends on w directly and on $(1-z')$ which depends on w; in both instances the dependence is monotonically increasing. R depends on $(1-z')$, which is decreasing in z. ■

The function $I(w)$ is the surplus generated if all of an employee's effort goes into individual production (i.e., $I(w) = pae^*(w) - C(e^*(w))$). This function is maximized at p, because with risk-neutral employees, the first best level of total effort is attainable by paying employees the full value of their output, p ("selling" the enterprise to the employee). Anything less leads to an inefficient level of total effort.

The function $B(w,z)$ is the increase in output (relative to I) that results from effort that is shifted to cooperative production. These benefits arise because the productivity of individual effort is less than that of cooperative effort (i.e., $a < 1$). When $z' = 0$ there is no cooperative production so B vanishes.

The function $R(w,z)$ is the payoff to employees, which include their reservation utility and the extra rents paid to defectors. By shifting effort from cooperative to individual production, defectors increase their measured performance above that of cooperators by hd. This results in expected wages that are hwd greater than that of cooperators for the $1-z'$ proportion of the workforce that are defectors. The rents are increasing in w because of a direct effect on the pay differential between cooperators and defectors and because of an indirect effect on the number of defectors. However, once w becomes sufficiently large ($wd \geq sz+1$), the extra rents are zero because with $z' = 0$ there are no cooperators and the firm can proceed to satisfy the individual rationality constraint of the defectors with equality.

The net result is that I is increasing in w, while $B - R$ is decreasing up to a point and then becomes constant (once $z' = 0$). Consequently the period payoff Π can have two local maxima, one in the interior, $(0,p)$, the other at p. The first maximum is due to the fact that I and $B - R$ move in opposite directions. The other is due to the fact that Π is increasing when $w > (sz+1)/d$ (since $B - R$ is

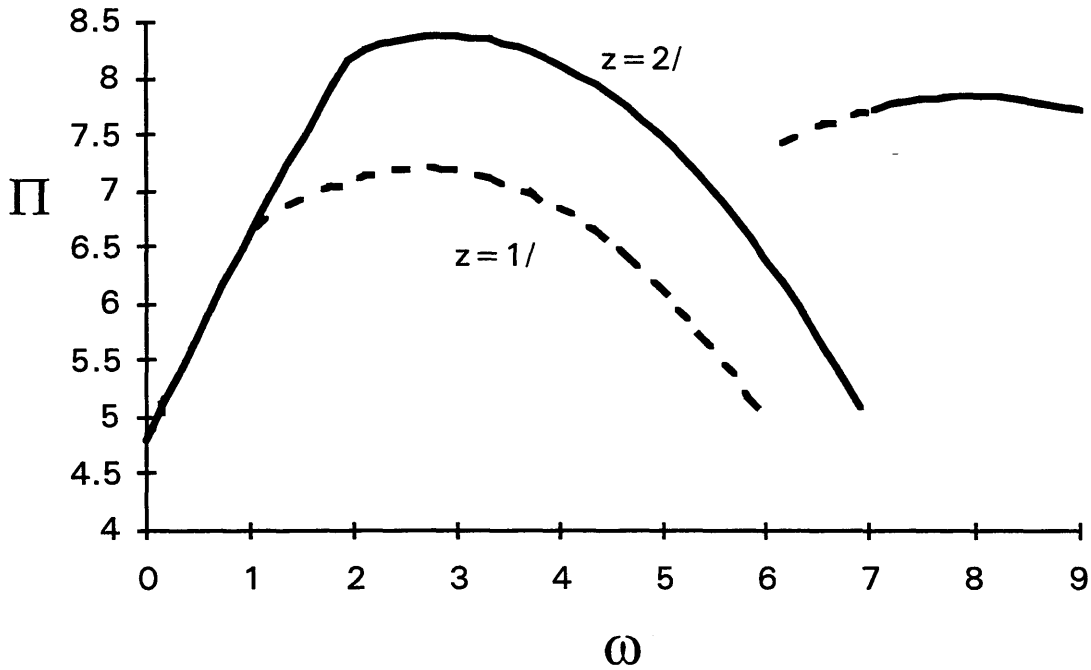

Figure 4.1: The One Period Profit

constant and I is increasing). As will be seen below, the existence of two maxima to Π is the key to the existence of multiple possible cultures and to the impact of history on the evolution of firms.

Figure 4.1 graphs the Π function for two values of z for the parameters $p = 8, a = .7, h = 2, c = 2$. Note the discontinuity caused by the elimination of the rents when z' goes to zero.

5. A Formalization of Corporate Culture

Our firm has a **corporate culture**. The model contains the basic elements of a culture: Employees have values (their sense of guilt), a range of possible behaviors (how cooperative they are) and socializing experiences within the organization (how much cooperation they receive from others). Figure 5.1 shows how these elements, plus the incentive intensity, interrelate to form a dynamic process.

The level of cooperative behavior by individuals e_C shapes the experiences of the workforce (i.e., $z = E[e_C]/h$). These experiences give rise to a distribution over values (i.e., $g \sim U[sz, sz+1]$) through the workings of social norms and reciprocity. These values determine the level of cooperation in the next period

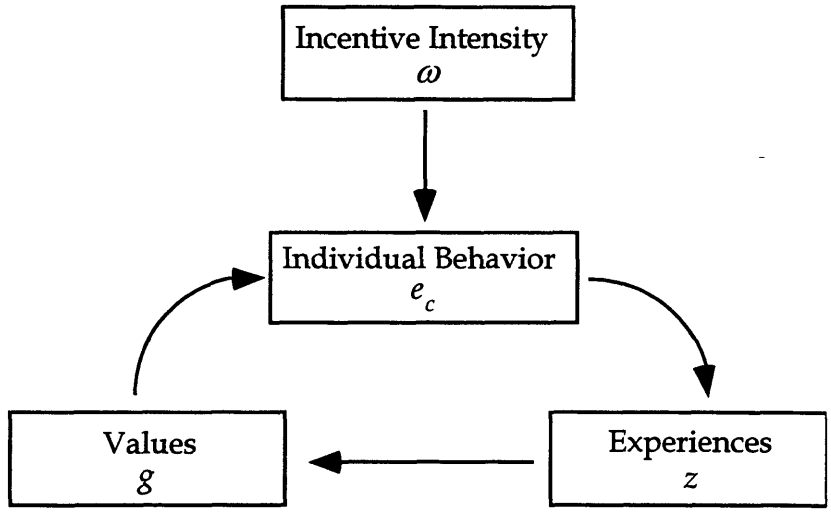

Figure 5.1: Culture as a Dynamic Process

(i.e., employees with $g > wd$ are cooperators), and the process continues. The cooperativeness of the firm's culture is summarized by the proportion of employees who are cooperators, z.

Our employees are not blindly controlled by social pressures. Rather, they trade-off their desire to behave in a socially responsible manner with the gains from pursuing their own self interests. That is, they trade-off the benefit of cooperating, avoiding feelings of guilt, against the cost, not focusing on individual production, which is more effective at raising an employee's measured output. Hence, the incentive intensity, which links measured output to employees' welfare, affects the extent of cooperative behavior and the evolution of the firm's culture.

In summary, a firm's culture in any period z_t depends on the prior culture z_{t-1} and the current incentive intensity w_t. This is just the relationship established in Lemma 2, $z_t = f(w_t, z_{t-1})$. An important consequence of having a corporate culture is that the firm's optimal choice of incentive intensity must solve a dynamic optimization problem. That is, z_t is a state variable, w_t is a control variable, $\Pi(w_t, z_{t-1})$ is the one period payoff and $f(w_t, z_{t-1})$ is the transition function. Expressed as a sequence program (choosing an infinite sequence of w_t's), the problem is written as follows.

$$(P)\, V(z) = \max_{(w_t)} \{J(w_1, w_2, ...)\} = \max_{(w_t)} \left\{ \sum_{t=1}^{\infty} \delta^{t-1} \Pi(w_t, z_{t-1}) \right\}.$$
$$\text{s.t. } z_t = f(w_t, z_{t-1}) \text{ and } z_o = z.$$

The firm's problem is similar to a growth model. The firm must trade-off

current "consumption," (i.e. a high incentive intensity and high current output) with investment in its productive "capital" (i.e., low incentive intensity, low output but a more cooperative culture in the future). A highly cooperative culture is like capital in that it produces a higher level of cooperation for a given incentive intensity. We highlight the similarity between capital and our firm's culture in the following lemma.

Lemma 5 *The value of the firm, $V(z)$, is increasing in z.*
Proof: Let $z^1 < z^2$. Consider some sequence (w_t) which is optimal given $z_0 = z^1$. Note that $\Pi(w, z)$ and $f(w, z)$ are strictly increasing in z for $w \in (sz/d, (sz+1)/d)$ and they are independent of z otherwise. Hence, (w_t) generates at least as great a payoff for $z_0 = z^2$ as for $z_0 = z^1$. ■

6. Analysis of the Dynamic Program

There is a potential technical problem with the program (P). The per-period return function $\Pi(w, z)$ has an upward discontinuity where the level of cooperation falls to zero (at $w = (sz+1)/d$) because the rent term vanishes. Fortunately, our program is equivalent to one with a continuous per-period return function.

Lemma 6 *There is no loss to maximizing (P) over the domain $w_t \in [sz/d, (sz+1)/d] \cup \{p\}$.*
Proof: Let $(w) = (w_t)_{t=1}^{\infty}$ be a candidate for a solution to (P). Then if one of the w_t's is in the range $((sz+1)/d, p)$, we can replace it with p. This will give a higher payoff in the current period and will result in the same $z_t = 0$. Hence it will generate a higher value for the objective. Likewise, if one of the w_t's is in the range $[0, sz/d]$ we can replace it by sz/d. Consequently, we can maximize (P) over the domain $[sz/d, (sz+1)/d] \cup \{p\}$. ■

Lemma 7 *Let $\widetilde{\Pi}$ be a continuous function which equals Π on $[0, (sz+1)/d] \cup \{p\}$ and lies below Π on $((sz+1)/d, p)$. Then the set of solutions to (P) coincides with the set of solutions to (P) when $\widetilde{\Pi}$ replaces Π.*
Proof: Since $\Pi \geq \widetilde{\Pi}$, the value of the objective for any (w) is no less under Π than it is under $\widetilde{\Pi}$. Let us call the maximization program under $\widetilde{\Pi}$, $(\widetilde{P})$. Consider a solution (w^*) to (P). Then, by Lemma 6 we can assume that none of the w_t^*'s is in the interval $((sz+1)/d, p)$. Therefore, the value of the objective in $(\widetilde{P})$ at (w^*) is the same. So (w^*) is a solution to $(\widetilde{P})$. Conversely, assume that we have a solution, $(\widetilde{w}^*)$ to $(\widetilde{P})$. If it was not a solution to (P) then we could find another sequence, say (w), which makes the objective in (P) higher and, by Lemma 6,

none of the w_t's is in $((sz+1)/d, p)$. But then the value of the objective in $(\widetilde{P})$ is higher at (w) than it was at $(\widetilde{w}^*)$, contradicting the optimality of $(\widetilde{w}^*)$. ■

We might as well seek and characterize solutions to the program with $\widetilde{\Pi}$ since it has the same set of solutions as the original program. For economy of notation, let us continue to refer to the new program as (P) and the new period payoff as Π. Now the period payoff is bounded and continuous. Therefore, we can draw on standard results from dynamic programming (see Stokey and Lucas (1989), henceforth SLP, chapter 4). In addition, we can now establish the following.

Lemma 8 *There exists a solution to the program* (P).
Proof: (P) is written as the maximization of $J(w) = J(w_1, w_2, ...)$ over the domain $[0,p]^\infty$. This domain is compact under the topology of weak convergence (Tychonoff theorem). Also, if we let $J_0 = 2\max\Pi(w,z)/(1-\delta)$, where w ranges over $[0,p]$ and z over $[0,1]$, then $J_0 < \infty$ (by the boundedness of Π). Now let $w^n \to w^\infty$ weakly. Then, for any T,

$$|J(w^n) - J(w^\infty)| \leq \delta^T J_0 + \max_{1\leq t\leq T-1} \{|\Pi(w_t^n, z_t^n) - \Pi(w_t^\infty, z_t^\infty)|\}.$$

Since $J_0 < \infty$, we can choose a T large enough to make the first term less than $\varepsilon/2$. Then, given this T and the continuity of Π, we can choose an n large enough that the second term is also less than $\varepsilon/2$. Therefore J is continuous under the topology of weak convergence. So it must have a maximum over $[0,p]^\infty$. ■

We now transform the program (P) so that we are maximizing with respect to the next period z rather than this period w:

$$(P^0) \max_{(z)} J^0(z_1, z_2, ...) = \max_{(z)} \sum_{t=1}^{\infty} \delta^{t-1}\Pi^0(z_{t-1}, z_t),$$

where $\Pi^0(z,z') = \Pi(\omega(z,z'), z)$ and $\omega(z,z') = \dfrac{1+sz-z'}{d}$.
The two maximizations are equivalent because the decision variables are strictly monotonically related over the relevant domain $[sz/d, (sz+1)/d] \cup \{p\}$. The following properties of Π^0 will be useful in the sequel.

Lemma 9 *The function* $\Pi^0(z,z')$ *is uniformly continuous over* $[0,1]\times(0,1)$ *and continuously differentiable over* $(0,1)\times(0,1)$.
Proof: Substituting ω into $\Pi(w,z)$ we obtain:

$$\Pi^0(z,0) = \frac{a^2}{c}\frac{p^2}{2} + \overline{e}ap - \overline{u}$$

$$\Pi^0(z,z') = \frac{a^2}{c}(p - \frac{1+sz-z'}{2d})\frac{1+sz-z'}{d} + \overline{e}ap + h(1-a)pz' - \overline{u}, \text{ for } z' > 0.$$

Therefore, Π^0 has a single discontinuity at $z' = 0$ and it is quadratic in z and z' on $[0,1] \times (0,1]$. ■

We denote the policy correspondence of (P) by $\omega(z)$ and the policy correspondence of (P^0) by $\zeta(z)$. Both are nonempty by Lemma 8. Our next result is that ζ is montonic.

Lemma 10 *Consider* z_1, z_2, z_1', z_2', *so that* $z_i \in [0,1]$, $z_1 < z_2$, *and* $z_i' \in \zeta(z_i)$, $i = 1,2$. *Then* $z_1' \leq z_2'$ *and if* z_1 or $z_2 \in (0,1)$, *then* $z_1' < z_2'$.
Proof: Assume $z_2' < z_1'$. Then, since z_1' is optimal at z_1, we must have: $\Pi^0(z_1, z_1') + \delta V(z_1') \geq \Pi^0(z_1, z_2') + \delta V(z_2')$, or

$$\delta[V(z_1') - V(z_2')] \geq \Pi^0(z_1, z_2') - \Pi^0(z_1, z_1').$$

We will now show that $\Pi^0(z_1, z_2') - \Pi^0(z_1, z_1') > \Pi^0(z_2, z_2') - \Pi^0(z_2, z_1')$. This together with above inequality shows that z_2' cannot be optimal at z_2. There are two cases to consider.
(a) $z_2' = 0$. Then, from the formula for Π^0, $\Pi^0(z_1, z_2') = \Pi^0(z_2, z_2')$ and $\Pi^0(z_2, z_1') > \Pi^0(z_1, z_1')$. So the desired inequality is established.
(b) $z_2' > 0$. Then

$$\Pi^0(z_i, z_j') = -\overline{u} - h\omega(z_i, z_j')d(1 - z_j') + hp(1-a)z_j' + I(\omega(z_i, z_j')),\ i,j = 1,2.$$

Therefore:

$$\begin{aligned}\Pi^0(z_1, z_2') - \Pi^0(z_1, z_1') &= h\omega(z_1, z_1')d(1 - z_1') - h\omega(z_1, z_2')d(1 - z_2') \\ &\quad + hp\,(1-a)\,(z_2' - z_1') + I(\omega(z_1, z_2')) - I(\omega(z_1, z_1'))\end{aligned}$$

and

$$\begin{aligned}\Pi^0(z_2, z_2') - \Pi^0(z_2, z_1') &= h\omega(z_2, z_1')d(1 - z_1') - h\omega(z_2, z_2')d(1 - z_2') \\ &\quad + hp\,(1-a)\,(z_2' - z_1') + I(\omega(z_2, z_2')) - I(\omega(z_2, z_1')).\end{aligned}$$

The term, $hp\,(1-a)\,(z_2' - z_1')$ is common and hence it will cancel. So it suffices to show:
(i) $h\omega(z_1, z_1')d(1-z_1') - h\omega(z_1, z_2')d(1-z_2') \geq h\omega(z_2, z_1')d(1-z_1') - h\omega(z_2, z_2')d(1-z_2')$.
and

(ii) $I(\omega(z_1, z_2')) - I(\omega(z_1, z_1')) \geq I(\omega(z_2, z_2')) - I(\omega(z_2, z_1')$.
But (i) is equivalent to

$$hd(1 - z_2')[\omega(z_2, z_2') - \omega(z_1, z_2')] \geq hd(1 - z_1')[\omega(z_2, z_1') - \omega(z_1, z_1')],$$

which holds because $\omega(z_2, z_2') - \omega(z_1, z_2') = \omega(z_2, z_1') - \omega(z_1, z_1')$ (see formula for ω) and because $z_1' \geq z_2'$. (ii) holds because $I(\cdot)$ is quadratic, the differences $\omega(z_1, z_2') - \omega(z_1, z_1')$, $\omega(z_2, z_2') - \omega(z_2, z_1')$ are equal, and $\omega(z_1, z_2') < \omega(z_2, z_2')$.

We now show that $\zeta(\bullet)$ is *strictly* monotonic whenever z_1' or $z_2' \in (0, 1)$. Assume $z_1' \in (0, 1)$ and let $z_1'' \in \zeta(z_1')$. Then, $z_1' \in \arg\max_r\{\Pi^0(z_1, r) + \delta\Pi^0(r, z_1'') + \delta^2 V(z_1'')\}$. But z_1' is interior and, hence, must satisfy the first order condition $\Pi_2^0(z_1, z_1') + \delta\Pi_1^0(z_1', z_1'') = 0$, where Π_i^0 is the partial derivative of Π^0 with respect to the ith variable. Since $z_1 < z_2$, $\Pi_2^0(z_1, z_1') < \Pi_2^0(z_2, z_1')$ and, thus, $\Pi_2^0(z_2, z_1') + \delta\Pi_1^0(z_1', z_1'') > 0$. Therefore, we can find a $\widetilde{z}_1 > z_1'$ so that $\Pi^0(z_2, \widetilde{z}_1) + \delta\Pi^0(\widetilde{z}_1, z_1'') + \delta^2 V(z_1'') > \Pi^0(z_2, z_1') + \delta\Pi^0(z_1', z_1'') + \delta^2 V(z_1'')$. This shows that $z_1' \notin \zeta(z_2)$. And since weak monotonicity of ζ has already been established, we must have $z_2' > z_1'$. A similar argument works for $0 < z_2' < 1$. ■

With montonicity in ζ, there is "critical mass" $\overline{z}$ of cooperators in the workforce such that when $z_0 > \overline{z}$ the level of cooperation need not go to zero. In particular, we can define

$$\overline{z} = \inf\{z > 0 \mid z' \geq z \text{ for some } z' \in \zeta(z)\}.$$

If $z' < z$ for all $z' \in \zeta(z)$ and all $z \in [0, 1]$, we let $\overline{z} = 1$. We now characterize the dynamics depending on whether the initial level of cooperation z_0 is above or below $\overline{z}$. We say that z_s is a **steady state culture** if it satisfies $z_s \in \zeta(z_s)$.

Lemma 11 *(i) For any $z_0 < \overline{z}$ (if there are such z_0's), any optimal sequence $(z_t)_{t=0}^{\infty}$ converges to the steady state $z_s = 0$. (ii) At that steady state $\omega(z_s) = p$.*
Proof: (i) From Lemma 9 and the definition of $\overline{z}$, it follows that $z_{t+1} < z_t$, for all t. Therefore, the sequence (z) must converge to a limit z^* and, by the upper-hemi-continuity of ζ (see SLP theorem 4.6), $z^* \in \zeta(z^*)$. Therefore z^* is a steady-state. Assume $z^* > 0$. Then this contradicts the definition of $\overline{z}$ since $0 < z^* < \overline{z}$. Therefore $z^* = 0$.
(ii) Starting from $z_0 = 0$, the firm follows the optimal path $z_t = 0$, i.e., it does not invest in culture. But then it might as well maximize its static profit by choosing $w = p$. ■

Consider now $z_0 > \overline{z}$ (if there are such z_0's) and let $(z) = (z_t)_{t=0}^{\infty}$ be an optimal sequence starting at z_0. Then, by the definition of $\overline{z}$, $z_t \geq z_0 > 0$. So we can restrict the maximization program (P^0) to sequences (z) which satisfy this

constraint. But, under this constraint, the period payoff is strictly concave and continuously differentiable. Consequently we have.

Lemma 12 *Consider $z_0 > \overline{z}$. Then: (i) The sequence program (P^0) has the same value and the same set of maximizers as the following dynamic programming program (with the decision variable w and the state variable $z > \overline{z}$):*

$$V(z) = \max_{0 \leq w \leq p} \{\Pi(w, z) + \delta V(f(w, z))\}. \tag{6.1}$$

(ii) There exists a unique value function, V. V is strictly increasing and continuously differentiable, with the possible exception of $z^* \equiv Inf\{z \mid \zeta(z) = 1\}$. (iii) There exists a (single-valued) policy function, $\omega(z)$, which is increasing in z. (iv) The culture converges to some steady state $z_s \in [\overline{z}, 1]$.

Proof: Equivalence follows from Π bounded and continuous and all $z_t > \overline{z}$. Furthermore, since Π is strictly concave, so is V over $(\overline{z}, 1)$. Uniqueness of the maximizing wage sequence follows from the strict concavity of Π and from theorem 4.8 in SLP. The differentiability of Π and theorem 4.10 of SLP imply that V is continuously differentiable at any z at which $\zeta(z) < 1$. Assume $\zeta(z) = 1$ for some $z < 1$ and let $z^* \equiv Inf\{z \mid \zeta(z) = 1\}$. Then, by Lemma 10, $\zeta(z) = 1$ and $\omega(z) = sz/d$ for any $z > z^*$. Thus $V(z) = \Pi(sz/d, z) + \delta V(1)$. Therefore V is differentiable for all $z > z^*$ and all $\overline{z} < z < z^*$. So the only possible non-differentiability is at z^*. The strict monotonicity of V follows from the fact that $z > \overline{z} > 0$ which implies $\omega(z) \in (sz/d, (sz+1)/d)$, a range over which Π is strictly monotonic.

Consider now the maximization programs on the RHS of (5.1) at z_1 and z_2, $\overline{z} < z_1 < z_2 \leq 1$ and let $w_i^* = \omega(z_i)$. Then we must have $\frac{\partial^-}{\partial w}\{\Pi(w_1^*, z_1) + \delta V(f(w_1^*, z_1))\} \geq 0$, where ∂^- denotes the left-hand derivative. But, since $z_2 > z_1$, $\frac{\partial^-}{\partial w}\{\Pi(w_1^*, z_1) + \delta V(f(w_1^*, z_1))\} < \frac{\partial^-}{\partial w}\{\Pi(w_1^*, z_2) + \delta V(f(w_1^*, z_2))\}$. So, given that the objective is concave, the maximizer w_2^* must be $\geq w_1^*$. ■

As a consequence of the Lemma 12 we have:

Lemma 13 *There exists at most one positive steady state.*

Proof: Assume there are two, $0 < z^1 < z^2$. Then $z^i \geq \overline{z}$, otherwise $z^1 < \overline{z}$ and, by Lemma 11, the sequence $(z_t)_{t=0}^{\infty}$ for which $z_0 = z^1$ converges to 0–contrary to the assumption that z^1 is a positive steady state. Now, according to equation ω, the wages which sustain these steady states are $w^i = [1 - (1-s)z^i]/d$, so $w^1 > w^2$. However, by Lemma 11, $w^1 \leq w^2$. Since these inequalities cannot hold simultaneously, this contradicts the existence of 2 positive steady states. ■

When the culture converges to an interior steady state z_s, the evolution of culture is characterized by a second-order difference equation.

Lemma 15 *Consider an optimal sequence $(z_t)_{t=0}^{\infty}$ such that $z_t \in (0,1)$ for all t. Then*

$$[hcd^2s + a^2s]z_{t-1} - [2hcd^2 + a^2 + \delta a^2s^2]z_t + [\delta a^2s + \delta hcd^2s]z_{t+1} =$$

$$a^2(1-\delta s)(pd-1) + \delta hcd^2s - h(1-a)cd^2p - 2hcd^2. \tag{6.2}$$

Proof: A necessary condition for an optimal and interior z_t is:

$$\frac{\partial J^0(z_1, z_2, ...)}{\partial z_t} = \delta^{t-1}\Pi_2^0(z_{t-1}, z_t) + \delta^t\Pi_1^0(z_t, z_{t+1}) = 0.$$

Or, after cancellation of δ^{t-1},

$$\Pi_2^0(z_{t-1}, z_t) + \delta\Pi_1^0(z_t, z_{t+1}) = 0. \tag{6.3}$$

Computing the partial derivatives of Π^0 (see Lemma 9) we have:

$$\Pi_1^0 = \frac{s}{d}[\frac{a^2}{c}(p-\omega) - hd(1-z')],$$

$$\Pi_2^0 = (\frac{-1}{d})[\frac{a^2}{c}(p-\omega) - hd(1-z')] + h(1-a)p + h\omega d,$$

where $\omega = (1 + sz - z')/d$. Substituting this into (6.3) gives the desired equation. ■

7. The Results

Our results point towards differences in culture, incentive intensity and performance across firms, even when they use the same technology and employ similar workforces. These differences come about because of differences in initial conditions. For example, when firms are first established, one firm may be more successful than others at hiring cooperative workers. This initial advantage is then magnified over time, and the two firms end up with very different cultures and very different profitabilities. In this sense the model exhibits *history dependence.*

Proposition 1: *(i)For any set of parameter values, the firm's culture converges to a steady state z_s starting from any initial condition, z_0. (ii) There are either*

one or two steady states for a given set of parameters. (iii) If there are two steady states there exists a $\overline{z} \in [0,1]$ so that if $z_0 < \overline{z}$, the culture converges to a steady state with $z_s = 0$ and $w_s = p$, while if $z_0 > \overline{z}$, it converges to a steady state with $z_s \geq \overline{z}$ and $w_s < p$. (iv) If there are two steady states, the firm with $z_s > \overline{z}$ has higher profits.

Proof: The convergence results follow from Lemmas 11, 12 and 13. The possibility of multiple steady states is demonstrated by Figure 3. The profit ordering, for $z_s > \overline{z}$, follows from Lemma 12. ■

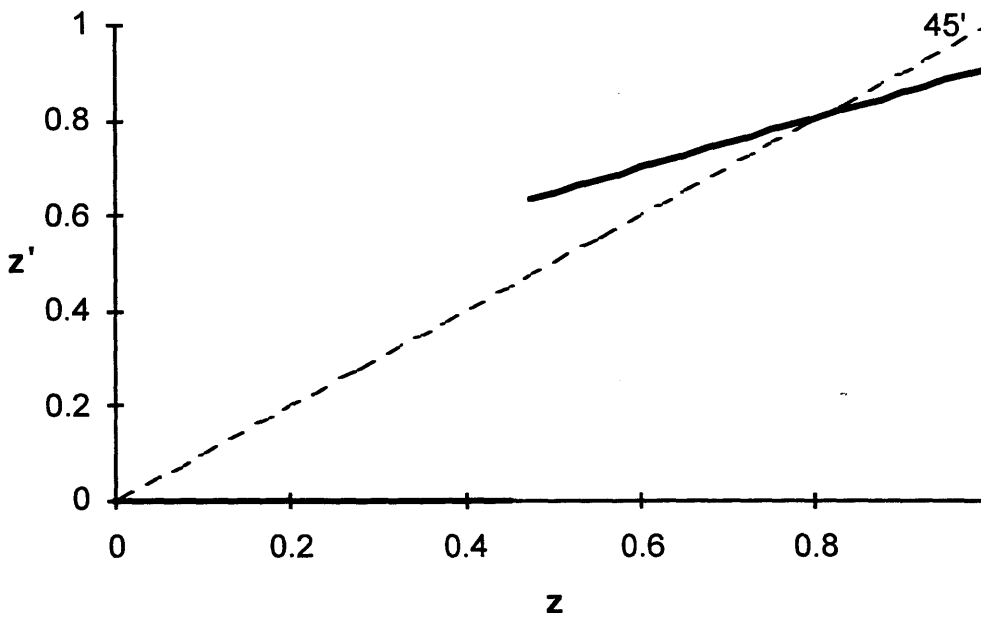

Figure 7.1: An Example of Multiple Steady States

Figure 7.1, which plots z' as a function of z for the parameters used in figure 2 and for $\delta = 0$, illustrates the existence of multiple steady states: For z_0 below $\overline{z} = 0.45$, the culture converges to $z_s = 0$. For z_0 above .45, the culture converges to z_s=0.82. When the $z - z'$ curve lies above the 45^0 line, z and w increase over time. The firm accumulates culture which enables it to put stronger incentive pressure on its workforce. Conversely, when the curve is below the 45^0 line, the firm milks its culture and reduces its incentive pressure over time. Therefore the convergence to a steady state is monotonic (these properties are shown formally

in Proposition 3). It is also possible to find parameter values for which there is a unique steady state with either $z_s = 0$ or $z_s > 0$.

The case of $\delta = 0$ is interesting because it shows the effect of complementarities between z and z', and these complementarities lead some firms into a "low corporate-culture trap". When z is low, it is costly to induce cooperation: Because workers are not cooperative to begin with, the firm must set a low incentive intensity which results in too little total effort. Conversely, when z is high, it is less costly to induce cooperation because workers are already cooperative. So a higher incentive intensity is still consistent with cooperation. Hence, it is possible that a firm with a highly cooperative culture "buys" a lot of cooperation—and thereby maintains its culture—while a firm with an uncooperative culture wants to buy little or no cooperation—thereby milking its culture. This helps explain why corporate culture is a hard-to imitate asset (Barney 1986): Once a firm has a low culture it is too costly for it to change. Figure 3 illustrates the existence of multiple steady states when $\delta = 0$.

When the firm is not myopic, $\delta > 0$, an additional, dynamic trade-off comes into play. The optimal incentive intensity balances the gains from increased total effort and profits in the present period against the cost of lower cooperation and lower profits in future periods. In figure 3 this is reflected in an upward shift of the $z - z'$ curve. This implies that the positive steady state is larger and its basin of attraction is bigger (see also Proposition 2). A far sighted firm is willing to suffer lower profits in the short run while it builds cooperative culture for the future.

If there are multiple steady states, the firm with the more cooperative culture is the more profitable; it also has lower incentive intensity. The higher profit arises because high cooperation enable large total effort *and* a large fraction of it devoted to cooperative production, both of which increase the firm's profit. On the other hand, to maintain the high culture the firm must reduce the incentive intensity. Consequently, if we consider a cross section of firms in a similar industry, the model predicts negative correlation between profitability and incentive intensity.

Let us examine now the effect of exogenous parameters in our model. To that end we determine the comparative statics of the interior steady state.

Proposition 2: *When it exists the interior steady state is*

$$z_s = \frac{hcd^2[2 + (1-a)p - \delta s] - a^2(1-\delta s)(pd-1)}{hcd^2[2 - s(1+\delta)] + a^2(1-s)(1-\delta s)},$$

$$w_s = \frac{hcd^2[s(1-\delta s) - p(1-a)(1-s)]}{hcd^3[2 - s(1+\delta)] + a^2 d(1-s)(1-\delta s)}.$$

z_s is increasing in c, h and δ, and is decreasing in p. The steady state incentive intensity moves in the opposite direction.

Proof: To find the expression for z_s, set $z_{t-1} = z_t = z_{t+1} = z_s$ in equation (6.2) and solve for z_s. To find the expression for w_s use the relationship $z_s = f(w, z_s)$. To differentiate z_s with respect to h, form $z_s = \dfrac{Ah+B}{Ch+D}$. Then $sign\{\partial z_s/\partial h\} = sign\{A(Ch+D) - C(Ah+B)\}$. This is positive since $Ch+D > Ah+B$ whenever $z_s < 1$, and $A = 2+(1-a)p-\delta s > 2-(1+s)\delta = C$. Likewise $\partial z_s/\partial c > 0$ since c and h are equivalent in the expression for z_s. To differentiate z_s with respect to p, form $z_s = \dfrac{Ap+B}{D}$. Since $B > D$, it must be that $A < 0$ whenever $z_s < 1$. To differentiate z_s with respect to δ form $z_s = \dfrac{A+B\delta}{C+D\delta}$. Then $sign\{\partial z_s/\partial \delta\} = sign\{B(C+D\delta) - D(A+B\delta)\}$. This is positive since $C+D\delta > A+B\delta$ and $B = -shcd^2 + sa^2(pd-1) > -shcd^2 - s(1-s)a^2 = D$. ■

The intuition behind the comparative static results for w_s is straightforward: As p increases, the firm wants more total effort, which calls for a higher w. As c increases, incentive intensity is less effective at increasing total effort. Hence, the firm chooses a smaller w. As h increases, cooperative production is more important and a cooperative culture is more valuable. Hence, the firm wants a higher z which requires a lower w. As δ increases, the firm puts a higher weight on future profits which, again, requires a higher z and a smaller w. The intuition behind the comparative static result for z_s is similar.[11]

The comparative static result with respect to δ can be interpreted as follows. Consider two firms, operating in two different countries. One firm faces a low δ, the other a high δ. Then the model predicts that the former would have a weaker corporate culture and a stronger incentive intensity than the former. This may help explain why American firms, which are said to be more myopic than Japanese firms, end up with a weaker loyalty of the workforce and with stronger incentive pressure.[12]

The next result shows the dynamics towards the interior steady state. In particular, it shows that the firm's culture and incentive intensity converge smoothly and monotonically towards the interior steady state (in case it exists).

[11] Interestingly, the comparative statics with respect to a could go either way. Initially, it might be thought that $\partial z_s/\partial a < 0$. However, when output from individual production increases (due to an increase in a), the firm might reshuffle its resources towards cooperative production, generating a larger z_s. (On the other hand, one can impose further restrictions on the model's parameters, ensuring that $\partial z_s/\partial a < 0$).

[12] Just as in one of the quotes from the introduction, an increase in the incentive intensity has short run benefits as total effort increases, but it can be harmful in the long-run as the level of cooperation in the organization decays over time.

Proposition 3: *Suppose there is an interior steady-state, $z_s \in (0,1)$ and $z_0 > \overline{z}$. Then, there exists a $\lambda \in (0,1)$ such that the firm's culture evolves as follows:*

$$z_{t+1} = (1-\lambda)z_s + \lambda z_t. \tag{7.1}$$

If $z_0 < z_s$, then $z_{t+1} > z_t$ and $w_{t+1} > w_t$; if $z_s < z_0$, then $z_{t+1} < z_t$ and $w_{t+1} < w_t$.

Proof: Starting from $z_0 > \overline{z}$, the sequence $(z_t)_{t=0}^{\infty}$ evolves according to the second-order difference equation (6.2). The general form of the solution is $z_t = z_s + A_1\lambda_1^t + A_2\lambda_2^t$, where λ_i's are the roots of the quadratic equation $g(x) = A + Bx + Cx^2$, with A, B and C being the coefficients of z_{t-1}, z_t and z_{t+1} in the expression in Lemma 14. It is readily verified that $g(0) > 0 > g(1)$ and $C > 0$. Therefore, there must be one root $\lambda_1 \in (0,1)$ and another root $\lambda_2 > 1$. But since we are looking for an interior steady state and the whole trajectory is in $(0,1)$, the explosive root λ_2 must have a coefficient $A_2 = 0$. The net result is that $z_t = z_s + A\lambda^t$ with $\lambda \in (0,1)$. Writing this formula for z_{t+1} and substituting for $A\lambda^t$ from z_t gives the result. The monotonicity of z and w is a consequence of equation (6.1). ■

As Proposition 3 suggests, the firm varies its incentive intensity over time. Hence the model predicts not only cross-sectional variation of the incentive system (as per Proposition 1), but also time-series variation. On the other hand, the time-series variation diminishes as the firm approaches the steady state. Nonetheless, if there is a random component to culture (e.g., as a result of workers turnover), the firm will vary its incentive intensity forever. This seems a plausible explanation for why firms keep varying their compensation systems even though their production technology and employees' preferences remain constant.

In discussing Proposition 1 we argued that the firms' optimal choice of incentive intensity plays a crucial role in creating multiple stable cultures. We demonstrate this formally by showing that for a *fixed* w and $s < 1$, which is what we have assumed, there is a *unique* steady state. Thus, varying the incentive intensity over time is necessary to bring about multiple steady states.

Proposition 4 *For $s < 1$ and a fixed w, there is a unique steady state z_s. For $s \geq 1$, and a fixed $w \in [1/d, s/d]$ there are multiple steady states.*

Proof: With a fixed incentive intensity w, a steady state satisfies $z_s = f(w, z_s)$. Since f is continuous with domain=range= $[0,1]$, there must be at least one steady state. Suppose now $s < 1$. Given the specific functional form for f given in Lemma 2, there are three possible steady states: 0, 1, and $(1-wd)/(1-s)$. We show uniqueness for each of four possible regions of w. First, if $wd \geq 1+s$, then $f(w,z) \equiv 0$, so $z = 0$ is the only steady state. Second, if $wd \in [1, 1+s)$,

then $f(w,z) < 1$ and $(1-wd)/(1-s) \leq 0$. So the unique steady state is $z = 0$. Third, if $wd \in (s,1)$, then $f(w,z) \in (0,1)$, and only the interior steady state, $(1-wd)/(1-s)$, is possible. Finally, if $wd \leq s$, then $f(w,z) > 0$ and $(1-wd)/(1-s) \geq 1$. So the only steady state is $z = 1$. (ii) Suppose $s \geq 1$ and $w \in [1/d, s/d]$. Then $f(w,0) = 0$ and $f(w,1) = 1$. Therefore both 0 and 1 are steady states. ■

The model with $s \geq 1$ and a fixed incentive intensity demonstrates the propensity of cultures to reproduce themselves, even without the optimizing behavior of firms. Lots of cooperative behavior begets employees who value cooperation and hence who reciprocate by behaving cooperatively themselves. Of course, the logic works as well for uncooperative behavior. If socialization processes are sufficiently strong, these positive and negative feedbacks are sufficient–on their own–to create multiple possible cultures.

8. Other Dimensions of Culture

The literature on organizational behavior differentiates between two types of motivation: those that are **extrinsic** to the task (e.g. incentive pay) and those that are **intrinsic** to the task; see Deci (1975). A taste for cooperative tasks is just one of many possible intrinsic motivators.[13] For example, employees might be motivated to perform certain tasks by a taste for producing high quality output (i.e., "pride in a job well done"), a belief in the mission of the firm, an enjoyment of intellectual challenges or a desire to act in a professional manner. We now briefly consider whether our theory generalizes to other forms of intrinsic motivation.

That our theory might generalize is suggested by the following quote from a senior partner at a law firm. After describing his firm's culture as having "a pride of craftsmanship" and a "family feeling", he goes on to talk about the angst that is being caused by the need to get partners to put more effort into certain tasks.[14]

> [I]t's becoming clear to us that we have to give increasing attention to issues such as business getting, being productive, billing efficiency and similar matters. While we all agree on that, we're not agreed on how far to go. We don't want to sacrifice our professionalism, our quality of service or our sense of colleagueship. What brings this all to a head is concern over our partner compensation system.

[13] Social concerns have both intrinsic and extrinsic elements. Motivation based on fear of social sanction is extrinsic. Motivation based on a desire to help others is intrinsic.

[14] The following quote is from Harvard Business School Case # 9-495-037 "Brainard, Bennis & Farrel."

The partners are concerned that if they put too much emphasis in the compensation system on the newly important tasks (e.g., a high incentive intensity linked to business getting), they will undermine valuable elements of their firm's culture related *both* to cooperation ("family feeling" and "sense of colleagueship") and to other forms of intrinsic motivation ("a pride of craftsmanship," "professionalism" and "quality of service").

Our theory is built on a few critical assumptions: that cooperative tasks are not fully reflected in the performance measures, that employees have a taste for performing cooperative tasks and that the taste for cooperation is increasing in prior levels of cooperation. There are then two key conditions required to apply our theory to other forms of intrinsic motivation. First, does the intrinsic motivator lead employees to put effort into tasks for which performance measures are lacking? Second, is the level of intrinsic motivation history dependent?

Does intrinsic motivation for providing high quality output satisfies these two conditions? In many settings, measuring the quality of production is much more difficult than measuring the quantity of output. Hence, a taste for high-quality output can be a useful motivator, and the first key condition is met in many settings. What of the second condition? Clearly, generalized reciprocity does not act here to create history-dependent preferences as the beneficiary of quality output is the firm rather than other employees. Social norms can still operate, although one expects that they will be less powerful in regards to quality than with regard to cooperation as employees have less opportunity to learn about the quality of their co-workers production.

There is, however, a third important source of history-dependent preferences which we have not yet discussed. According to what psychologists call escalating commitment (see Baron and Kreps, forthcoming), the more people act according to some motivation, the stronger is that motivation. This psychological phenomena is familiar to economists as a source of the often encountered fallacy of sunk costs. With escalating commitment, a firm that puts incentive pressure on employees today will see them cut corners on quality and hence, over time, come to value quality less and less. Then, the second key premise of history dependent preferences is met and we expect that the sort of cultural dynamics that we identified with cooperation could occur with tastes for quality as well. We conclude that our theory extends to other forms of intrinsic motivation and hence to other dimensions of corporate culture.[15]

[15]Extending our theory to general intrinsic motivation elucidates another claim by the management consultant Kohn: "Contrary to the conventional wisdom, the use of rewards is not a response to the extrinsic orientation exhibited by many workers. Rather, incentives help create this focus on finacial consideration." A high incentive intensity not only undermines a taste for cooperation, but other intrinsic motivators as well.

9. Conclusion

As Milgrom and Roberts acknowledge at the beginning of *Economics, Organizations and Management*, "important features of many organizations can best be understood in terms of deliberate attempts to change preferences of individual participants" (p. 42). This paper offers one of the first formal attempts to show how business practices can be understood as attempts to change preferences. By considering endogenous preferences, we are able to propose an alternative to the standard risk-sharing theory of incentive intensity, to formalize the concept of corporate culture and to begin exploring the nature of culture-based performance differences across firms. We close with thoughts on future work.

We implicitly assumed that there was no turnover in the firm's workforce. If this is relaxed, the preferences of employees who leave and join the firm will be important for how the culture evolves. Hence, human resource policies that affect turnover and the screening of new hires can also be studied based on their effect on the evolution of a firm's corporate culture. We already have one interesting observation on this topic: Because defectors earn rents in firms with cooperative cultures, our theory suggest that agents with low inherent feelings of guilt (low γ) are especially attracted to firms with cooperative cultures. This suggest that firms with highly cooperative cultures must take special measures to screen out applicants with a low sense of social responsibility and to retain those with a high sense of social responsibility. Otherwise, they may see their culture eroded over time.

While there is some laboratory work that is consistent with history-dependent preferences, more study is needed. It is possible to recreate in the laboratory the choices made by employees in our model. Subjects could play a series of one shot PD's in which the incentive for defecting is independent of the opponent's play. To our knowledge, no one has looked at how an individual's past experience correlates with how she plays such games. One objective would be to explore the source of history-dependent preferences. We suggested two possibilities, generalized reciprocity and adherence to social norms. These effects can be separated (to some extent) by seeing how much personal history still matters when subjects are told the overall level of cooperation in the pool of subjects.

We assumed that an employee's opportunities for cooperating, as parameterized by h, were exogenously given. It is possible to have h endogenous. The extent of an employee's cooperative opportunities could depend on decisions made by both other employees and the firm. For example, the less cooperative the culture, the less likely employees are to bother asking each other for help. Similarly, firms with more cooperative cultures will be more willing to choose technologies where employees are highly interdependent and cooperation is important. We

hypothesize that such effects serve to reinforce the barriers to imitation that we have identified and to make changing cultures even more costly.

We were careful to construct a model without strategic interactions among workers. This is a simplification. Consider that many cooperative tasks may involve joint work by several employees. The output from such joint work may exhibit complementarities between individual efforts. Further, employees may have preferences not just over their own action, but also over how their action compares to the actions of their partners. If either of these effects are present, strategic interactions arise and employees expectations or beliefs about coworker behavior becomes important. (Beliefs are also cited as being part of the phenomenon of culture.) We hypothesize that strategic interactions make multiple stable cultures more likely to arise and make it more difficult for firms to change cultures.

References

[1] Aronson, E. (1972), *The Social Animal,* San Francisco: W.H. Freeman and Company.

[2] Athey, S., C. Avery and P. Zemsky (1994), "Mentoring, Discrimination and Diversity in Organizations," Working Paper, Stanford Graduate School of Business.

[3] Baron, J. and D. Kreps (forthcoming), *Human Resource Management for General Managers,* New-York: John Wiley and Sons.

[4] Barney, J. (1986), "Organizational Culture: Can it be a Source of Sustained Competitive Advantage," *Academy of Management Review* 11, 656-665.

[5] Berg, J., J.Dickhaut and K. McCabe (1995), "Trust, Reciprocity and Social History," *Games and Economic Behavior* 10, 122–142.

[6] Dawes, R. (1980),"Social Dilemmas," *Annual Review of Psychlogy* 31, 169–193.

[7] Deci, E., (1975) *Intrinsic Motivation,* New York/London: Plenum Press.

[8] Forsythe, R., J. Horowitz, N.E. Savin and M. Sefton (1994), "Fairness in Simple Bargaining Games," *Games and Economic Behavior* 6, 347–369.

[9] Hoffman, E., K. McCabe, K. Shachat and V.Smith (1994), "Preferences, Property Rights, and Anonimity in Bargaining Games," *Games and Economic Behavior* 7, 346–380.

[10] Holmstrom, B. and P. Milgrom (1991), "Multi-Task Principal-Agent Analysis: Incentive Contracts, Asset Ownership and Job Design," *Journal of Law, Economics and Organization* 7, 24–52.

[11] Kreps D. (1990), "Corporate Culture and Economic Theory," pages 90-143 in *Perspectives on Positive Political Economy,* Edited by James E. Alt and Kenneth A. Shepsle, Cambridge MA: Cambridge University Press.

[12] Lazear E. (1995), "Corporate Culture and Diffusion of Values," pages 89-133 in *Trends in Business Organization: Do Participation and Cooperation Increase Competitiveness?,* Edited by Horst Siebert, Tubrigen: Mohr Publishing Company.

[13] Milgrom P. and J. Roberts (1992), *Econmics, Organization and Management,* Englewood Cliffs, NJ: Prentice Hall.

[14] O'Reilly, C. (1989), "Corporations, Cultures and Committments: Motivations and Social Controls in Organizations," *California Management Review* 31, 9–25.

[15] Rotemberg, J.J., (1994) "Human Relations in the Workplace," *Journal of Political Economy* 102, 684-717.

[16] Stokey, N., R. Lucas and E. Prescott (1989), *Recursive Methods in Economic Dynmics*, Cambridge, MA: Harvard University Press.