Economics & Corporate Culture*

Benjamin E. Hermalin†
University of California, Berkeley
February 2000

1 Introduction

Most non-economists would question whether economists know anything about culture. Indeed, given some faculty meetings I've been through, I imagine they might go so far as to say economists are uncultured. Although the second sentiment probably does go too far, there is certainly merit to the first. In particular, with a few exceptions (notably Kreps, 1990; Crémer, 1993; Lazear, 1995; Hodgson, 1996), economists have ignored the issue of corporate culture in their studies of firms and other organizations.[1] There are many reasons for this lack of attention: culture is not relevant in most economic modeling; culture is not rational (or at least not obviously so) and, hence, does not fit well with the rational-agent methodology of neoclassical economics; and culture is difficult to define or measure, making it hard to use or control for in econometric analyses.

By writing this chapter, I'm agreeing with the proposition that corporate culture is worthy of study by economists and is amenable to our methods. Worthy because corporate culture is an important determinant of firms' capabilities and performance. Moreover, it both complements and substitutes for many of the other governance structures that economists have long studied. By amenable, I don't necessarily believe a complete *economic* theory of corporate culture is feasible. But even so, economics can contribute to a better understanding of

*The author thanks Jennifer Chatman, Jacques Crémer, David Kreps, and participants at the Berkeley Faculty Colloquium for comments on an earlier draft. The author also acknowledges the financial support provided by the Willis H. Booth Chair in Banking & Finance, Cornell's Johnson Graduate School of Management, and the NSF under grant SBR-9616675.

†**Contact Information:**
Professor Benjamin E. Hermalin
University of California
Walter A. Haas School of Business
545 Student Services Bldg. #1900
Berkeley, CA 94720-1900

Phone: (510) 642-7575; E-mail: `hermalin@haas.berkeley.edu`

[1]Although, as Crémer (1993) notes, at points, the discussion in Arrow's (1974) book on organizations contains ideas that clearly foreshadow what would later be called corporate culture. Hodgson (1996) suggests that the "old" (pre-Coase, 1937, and Williamson, 1975) institutional tradition in economics (e.g., the work of Thorstein Veblen) could also be seen as sympathetic to current notions of corporate culture.

corporate culture by shedding light on specific facets of corporate culture that are less amenable to analysis by other social sciences.

This chapter begins by reviewing, synthesizing, and commenting on earlier work by economists on corporate culture, with particular attention spent on Kreps's famous article. The second half of the chapter is spent discussing how certain insights from other economic analyses of organizations can complement our understanding of corporate culture.[2]

2 Kreps (1990) Reconsidered

Kreps offers two reasons for economists to consider corporate culture. First, understanding culture—and organizations more generally—is necessary for understanding how firms implement strategy:

> The actual purpose of the firm *qua* organization is not considered [by textbook economics]. This is rather strange, for if one has an economic mind-set, one must believe that the firm itself performs some economic (efficiency-promoting) function. From there it is a short step to consider as part, perhaps the largest part, of successful strategy those actions designed to increase the firm's organizational efficiency. (From the Introduction.)

Kreps's second reason is his belief that economics have now developed the theoretical tools to study culture and he wants "to present the outlines of the theory that is developing," while encouraging his readers to develop it further.

2.1 The Kreps (1990) Model

Kreps's analysis of corporate culture is built from the following ingredients:

- *Formal* contracts are costly or defective in many situations. As Williamson (1975) and others have observed, formal contracting can be costly: There are bargaining costs, costs associated with monitoring the agreement, and costs associated with enforcing the agreement. In addition, formal contracts contingent on the appropriate variables can be infeasible: states or actions upon which the ideal contract would be contingent are not verifiable[3] or, somewhat equivalently, too difficult to specify in advance. That is, as economists would put it, only *incomplete* contracts could be feasible in many situations of interest.

- Firms are repeat players. Contracts are one way to induce cooperation from parties who would otherwise have incentives not to cooperate. Repeated games are another way: One player's deviation from cooperating

[2]The reader interested in the more general topic of the economics of organization (with, however, little to no discussion of corporate culture) would do well to consider the surveys by Gibbons (1998, 1999) and MacLeod (1995). MacLeod's survey is the most technical of the three.

[3]Following the convention in economics, a variable is *verifiable* if its value can be learned by a judge or other outside party called upon to adjudicate a dispute.

today can be punished by other players refusing to cooperate in the future. Provided the one player's future benefits from all cooperating in the future outweigh the gains today from his not cooperating, this punishment will deter non-cooperative behavior.[4] Critical to exploiting this benefit from repeated games is that one player, at least, be long-lived; that is, able to play in multiple periods. Firms satisfy this requirement and, typically, so do their senior officers.

- In many situations, inducing cooperation through repeated play is cheaper than inducing it contractually. For example, it would seem cheaper to have an oral agreement to cooperate that's enforced by repeated play than to have lawyers draw up a formal agreement that obtains the same result. In other situations, desirable outcomes can be supported in repeated games that cannot be supported contractually. In particular, it could well be that the variables upon which the ideal contract would be written are observable to the parties—they know whether each of them is cooperating and who defected if cooperation breaks down—but not verifiable. In this case, repeated play can achieve what formal contracting cannot.

- **Multiple equilibria.** Many games, included repeated ones, have multiple equilibria (i.e., stable outcomes). The players may need some means of coordinating on which equilibrium they are to play.

- **Unforeseen contingencies.** Not all contingencies can be foreseen.[5] In a world in which there are unforeseen contingencies, parties may need to trust each other to do the right thing should such a contingency arise. Unforeseen contingencies are often given as another reason for contractual incompleteness.[6]

The first three of these ingredients have a role to play in making sense of corporate culture in Kreps's model, but do not in themselves explain it. After all,

[4]Among economists, this benefit from repeated interaction is often cited as the "folk theorem." Among non-economists, it is often associated with Axelrod (1984), who found that the best strategy in a repeated prisoners' dilemma game was the "tit-for-tat" strategy that rewarded cooperation and punished non-cooperation.

[5]Despite superficial similarity, the notion of unforeseen contingencies is distinct from the idea that some contingencies are difficult to specify in advance: The latter idea refers to contingencies that the parties can foresee but have difficulty specifying, while the former refers to contingencies that they don't even foresee.

[6]This view, however, is not as straightforward as it might seem: As long as the parties to a contract recognize that there could be unforeseen contingencies, they can still write a "complete" contract in the sense that they can insert a "none-of-the-above clause" into the contract; that is, a clause of the form, "should some contingency not mentioned above occur, then we agree to do ..." Hence, to have what Hermalin and Katz (1993b) call a *literally* incomplete contract requires assuming the parties are boundedly boundedly rational—they can't foresee that they might have failed to foresee certain contingencies. On the other hand, even if they are merely boundedly rational, a none-of-the-above clause might be problematic because it fails to makes distinctions among different unforeseen contingencies. For this reason, culture, which could make appropriate distinctions, could be necessary. See the introduction to Hermalin and Katz for a more in-depth discussion of these issues.

		Junior	
		A	B
Senior	A	$3,2$	$0,0$
	B	$0,0$	$2,3$

(a)

		Junior	
		A	B
Senior	A	$3,2$	$4,-4$
	B	$-4,4$	$2,3$

(b)

Figure 1: Coordination games between junior and senior personnel.

there is a large body of economics that makes use of these ingredients to explain a wide variety of organizational phenomena, none of which could be described as culture.[7] It is the final two elements—multiple equilibria and unforeseen contingencies—that "introduce" culture. Although Kreps seems at times to suggest that both these elements are needed, a more careful reading of his work suggests that each is sufficient, sometimes in combination with the first three elements, to provide insights into culture.

2.1.1 Multiple Equilibria

To illustrate how multiple equilibria relate to culture in Kreps's model consider Figure 1(a),[8] which illustrates a game to be played in some firm between a senior person and a junior person. Each player has a choice of two actions (strategies), A or B. There are, thus, four possible outcomes corresponding to the four possible pairings of actions. The payoffs to the players corresponding to these four outcomes are shown as the numbered pairs in the cells of the table in Figure 1(a). The first number in each pair is the payoff to the senior person should that outcome occur and the second number is the payoff to the junior person. Hence, for example, if the senior person chose A and the junior person also chose A, then the senior person would get 3 and the junior person would get 2.

This game has two pure-strategy Nash equilibria: In one, each player takes action A; in the other, each player takes action B.[9] What standard game theory does not tell us, however, is which of these two equilibria we should expect to see played.[10] Culture, on the other hand, might. Suppose that it is understood, as matter of culture, that juniors are to defer to seniors (readers at

[7]Broadly, these come under the heading of *implicit contracting*. See Section 2 of Hart and Holmström (1987) for a survey. There has also been a large literature—know as the *incomplete contracts literature*—that has explored just the first of Kreps's ingredients, parties' *in*ability to write ideal (i.e., complete) contracts (see §3 of Hart and Holmström, 1987, or Hart, 1995 for partial surveys).

[8]The game-theory *cognoscenti* will recognize this as the "battle of the sexes" game.

[9]Recall that an outcome is a Nash equilibrium if neither player would *unilaterally* wish to change his or her strategy given the strategy his or her opponent is to play. Hence, for instance, if both are to play A, then neither would unilaterally want to change to B because to do so would reduce his or her payoff from a positive amount (3 or 2) to 0. Likewise if both are to play B, a unilateral deviation would mean a reduction in payoff.

[10]As a technical matter, there is also a third Nash equilibrium in mixed strategies. There is, however, no need to explicitly consider this equilibrium in what follows.

universities—particularly in Europe—can no doubt relate to this norm). Consequently, juniors and seniors expect that the A–A equilibrium will be played since it is the better equilibrium for the senior player. Put somewhat differently, absent culture, nothing allows us—and, thus, the players themselves—to predict which equilibrium will be played. Neither we, nor the players, can, therefore, know which equilibrium is to be played. For the players, this is something of a disaster, since nothing then prevents one player choosing A and the other B; that is, they risk getting nothing. A strong cultural norm of deferring to seniors, however, eliminates this risk. In other words, culture increases predictability of others' actions.

Observe that only one of Kreps's ingredients has been used, multiple equilibria. The other ingredients can be added in to yield a more flavorful story. For instance, instead of "leaving it open" as to which equilibrium is played, the players could contract in advance to play the A–A equilibrium. A contract could, for example, stipulate that a player who plays B must pay the other player 4. This transforms the game into the one in Figure 1(b). This modified game has just a single equilibrium in which each player chooses A.[11] This contract, therefore, duplicates what the cultural norm achieves. If, however, this contract is costly to write or enforce, then it could be worse than relying on the cultural norm. Alternatively, the players' actions and payoffs could be unverifiable, which renders a contract such as this infeasible.[12] Unforeseen contingencies can also be tossed into the stew: Consider Figure 2. Provided $S > s > 0$ and $J > j > 0$, this generalizes the game of Figure 1(a). Imagine, now, that there are many such coordination games that junior and senior personnel will be called to play. What A will be and what B will be is not known *ex ante*, nor are the values of S, s, J, and j known. A player may not even know in advance if he will be the junior or senior player. Moreover, we can readily imagine that the potential players cannot even foresee the details of all possible coordination games they may face; that is, the details could be unforeseen contingencies. In such a situation, contracting is probably infeasible: The cost of verifying every time the necessary details of the game—in particular is it a coordination game fitting the model of Figure 2—could be prohibitive. Indeed, it is easy to imagine that the relevant variables aren't verifiable at any cost. Yet, if there's a cultural norm in place of deferring to seniors, then the players are nevertheless assured of avoiding disastrous outcomes in which one plays A and the other plays B.

This discussion suggests that culture can avert disastrous outcomes. This is not the same, however, as saying that culture leads to optimal outcomes. For the game of Figure 1(a), the cultural norm does lead to an optimal outcome: The A–A outcome in Figure 1(a) maximizes joint payoffs—5 is the largest possible

[11] It is readily seen that A is a *dominant* strategy for each player: No matter what she expects her opponent to play, she does better to play A.

[12] Since this contract also requires one party to pay the other a penalty, it could also be infeasible as a matter of law—a court could refuse to enforce it in the case of a A–B or B–A outcome, citing the *penalty doctrine* or the law's abhorrence of forfeiture (see, e.g., the Restatement (Second) of Contracts, §356 [St. Paul: American Law Institute Publishers, 1981]).

		Junior A	Junior B
Senior A		S, j	$0, 0$
Senior B		$0, 0$	s, J

Figure 2: General coordination game between junior and senior personnel.

sum—so it is, therefore, Pareto efficient. For the more general game (Figure 2), the story is more complicated. If we assume that transfers between the players are *not* feasible, then the A–A outcome is still Pareto efficient—a move to any other outcome would make the senior player worse off. If, however, transfers are feasible and $S + j < s + J$, then the players would be better off playing the B–B equilibrium: The junior player can make a transfer, t, to the senior player such that $s + t > S$ and $J - t > j$. Therefore, whenever $S + j < s + J$, the defer-to-seniors cultural norm does *not* lead to an optimal outcome (at least if transfers are feasible).

In a world in which transfers are feasible, one might imagine an extension of the defer-to-seniors norm that leads to optimality: Both play A unless $S + j < s + J$. In that case, both play B and the junior player transfers

$$t^* = \frac{S + J - s - j}{2}$$

to the senior player (a little algebra reveals that $s + t^* > S$ and $J - t^* > j$).[13] Call this the extended-deference norm. The problem with this norm, unlike the basic defer-to-seniors norm, is that it is *no longer rational for one of the players to abide by it.* In particular, why, having played the B–B outcome, would the junior player make herself worse off by actually transferring t^* to the senior player? She would do better to pocket the whole J.

This is where, *possibly*, repeated games can come to the rescue. It is also where we can make this more of a story about a *corporate* cultural norm. To make the discussion slightly more concrete, suppose that, with probability p, the game to be played is the "A-game" in Figure 3, while, with probability $1 - p$, it's the "B-game." Assume $S > 2$. Observe that, in the A-game, the welfare-maximizing equilibrium is, thus, the A–A equilibrium, while, in the B-game, it's the B–B equilibrium. Suppose that a given player lives two periods. In the first period of his life, he's a junior. In the second (last) period of his life, he's a senior.[14] Suppose the cultural norm is the extended-deference norm (here $t^* = 5$), but now suppose there is a punishment imposed on a junior who violates it. Specifically, when a norm-violating junior becomes a senior, the next generation junior plays B regardless of which game is realized. Moreover,

[13] Actually, any t such that $S - s < t < J - j$ would do—I chose the mid-point simply to be concrete.

[14] Although distinct from Crémer (1986) in formulation and interpretation, it is worth pointing out that Crémer also explores an overlapping generations game in the context of an organization.

		Junior	
		A	B
Senior	A	$S, 1$	$0, 0$
	B	$0, 0$	$1, 2$

A-game

		Junior	
		A	B
Senior	A	$2, 1$	$0, 0$
	B	$0, 0$	$1, 10$

B-game

Figure 3: Possible coordination games between junior and senior personnel.

this new junior makes no transfers to the senior. Provided a junior has either followed the extended-deference norm in playing against a norm-adhering senior or punished a norm-defying senior, he expects that the extended-deference norm will apply when *he's* a senior. Otherwise, he expects to be punished. Consider a junior playing against a norm-adhering senior. Clearly, there is no benefit to him from playing B in the A-game (he'll get 0 instead of 1). There's also no benefit to him from playing A in the B-game (he'll get 0 instead of at least 5). The only question is whether he will transfer 5 in the B-game. Suppose he doesn't, then he gets 10 today. Next period, he anticipates that he'll play against a B-playing junior (I'll verify in a moment that this junior will want to play B). Hence, his best response is to play B as well, which yields him a payoff of 1 regardless of which game occurs. His total payoff from violating the norm is 11.[15] Now suppose he adheres to the norm when he's a junior in the B-game. Then, today, he gets 5. Next period, he anticipates that he'll play against a norm-adhering junior (again, I'll verify this in a moment). Hence, with probability p, he'll get S and, with probability $1-p$, he'll get 6 $(= 1 + t^*)$. His total expected payoff is $5 + pS + 6(1-p)$. This is greater than 11 if $S > 6$; that is, when $S > 6$, he'll want to adhere to the norm. Note these calculations apply to any junior, so if he's adhered to the norm, he can expect the next generation's junior to do so as well. The only thing left to check is that a junior who faces a norm-defying senior will punish. Observe that, if the senior believes she will be punished, she will rationally play B, so the junior gets 0 today if he plays A. He will also be punished tomorrow, which means he gets 1 tomorrow. His total for both periods if he *doesn't* punish when he should is 1. If, instead, he punishes (plays B) when he should, then his expected payoff today is $2p + 10(1-p)$ and his expected payoff tomorrow is $pS + 6(1-p)$—a total of $(2+S)p + 16(1-p)$, which exceeds 1. Hence, he does better to punish when he's supposed to: He will, indeed, punish a norm-defying senior. Putting all this together: If $S > 6$, then the extended-deference norm can be supported as an equilibrium of a repeated game.

Some comments on this last analysis. First, repetition *won't* save the day if $S < 6$; that is, repetition won't always help (a point to which Kreps also alludes). Second, note that the above analysis depends critically on each generation's junior knowing precisely how the previous generation's junior behaved. This

[15]The technically sophisticated reader will have noticed that I've not discounted the second period payoff. Given the finite life of each player there is no need to do so; qualitatively similar results would hold if I introduced discounting.

is where, arguably, the "corporate" comes into corporate culture—corporations are well-known as repositories of information (Crémer, 1993; Hodgson, 1996). Hence, for instance, this generation's junior may hear others in the corporation gossip about what a jerk such and such a senior was.[16] As Kreps points out, even if information about past deviations isn't perfectly accurate (e.g., gossip is sometimes inaccurate or sometimes fails to inform), a benefit from repetition can still be had provided this information is reasonably accurate.

The analysis so far has illustrated the following: The need to choose among multiple equilibria can, in itself, provide a role for culture. The other ingredients are not necessary, although they certainly add some flavor to the analysis. At times Kreps, himself, hints that selecting among multiple equilibria is all that's needed to introduce culture into economic models (for example, when he discusses why American students will settle on particular equilibria in a game that calls upon them to divide a list of American cities—American culture, in particular a sense for the importance of the Mississippi river or the Mason-Dixon line as dividers, causes the players to settle on particular equilibria). At other times, however, he seems to suggest that the other ingredients are equally important. As I've tried to show here, that would be equivalent to confusing the beef with its seasonings.

Before concluding this section, it's worth pointing out that there's another literature in economics that appeals to culture to choose among multiple equilibria. Simple casual empiricism reveals that even within the same industry different firms behave differently. One explanation is that competition among these firms leads to asymmetric equilibria—it's not an equilibrium to behave identically. This is the approach I took in Hermalin (1994). Another approach, more relevant to the current discussion, is the one pursued by Okuno-Fujiwara (1994) and Morita (1998): They note, like many other observers, that American firms and Japanese firms exhibit different internal behaviors even when in the same industry. These authors develop models of firm organization that have multiple equilibria and argue that one equilibrium is consistent with the behavior of American firms and another is consistent with the behavior of Japanese firms. Related to the discussion here, they suggest that it is cultural differences between the U.S. and Japan that has led the firms in these two countries to settle on different equilibria with respect to their organization. Admittedly, these authors are relying on *national* culture, which is not the same as *corporate* culture. On the other hand, certainly national, regional, or professional cultures must influence corporate culture.[17] That is, measured differences in the latter could be due, in part, to differences in the former. In addition, the organizational differences that national or regional cultural differences induce could appear to be the consequence of differences in corporate culture.

[16]Indeed, a prominent phenomenon in any society is its eagerness to identify (*label*) its deviants (see, e.g., Erikson, 1966, or, somewhat less on point, Goffman, 1963).

[17]See Hofstede et al. (1990) for evidence that national culture affects firms and Chatman and Jehn (1994) for evidence that industrial culture (characteristics) affect firms. Vandello and Cohen (1999) provides evidence on regional variations in culture norms across regions of the United States.

		Employee	
		Trust	Don't trust
Boss	Treat fairly	7, 5	2, 3
	Exploit	9, 0	4, 1

Figure 4: Game between boss and employee.

This discussion raises another issue: Where does a corporate culture come from? Like any "input" used by the firm, corporate culture would seem subject to the make-or-buy decision. Does a firm essentially rely on the prevailing (i.e., national, regional, or professional) cultural norms or does it craft its own? Even without knowing how cultures are made, it seems reasonable to expect that crafting your own culture is costly and, therefore, not worth it for some organizations—they, instead, will "buy" the prevailing culture (i.e., relying on the societal or professional norms recruits bring with them rather than socializing them). Hence, although these organizations may be seen as having weak cultures—i.e., not distinct from the prevailing culture—their culture could still matter, particularly for interregional or international comparisons as suggested by the work of Okuno-Fujiwara (1994) and Morita (1998).[18]

2.1.2 Unforeseen Contingencies

As interesting as multiple equilibria are to the theory Kreps sketches, much of what he has to say works without any reference to multiple equilibria. To illustrate this, consider the game in Figure 4:[19] This is a game between a boss and an employee. The boss can treat the employee fairly or exploit him. The employee can play in a trusting way or in a not-trusting way. The latter strategy offers the employee some protection against exploitation, but is harmful to him if the boss plays fairly. Observe that *exploit* is a dominant strategy for the boss—against either strategy of the employee, her payoff is greater if she exploits than if she treats fairly. She can, therefore, be expected to exploit. The employee's best response to being exploited is *don't trust*. Hence, the equilibrium outcome is the bottom-right cell. Observe that this outcome is not optimal: Both players would be better off in the top-left cell. Unhappily, as we've just seen, if the game in Figure 4 is played just once, then this better outcome is *un*attainable.

As discussed above, repeated games can come to the rescue. In particular, suppose that the game is repeated next period with probability β, $0 < \beta < 1$. Imagine, too, that this probability remains constant.[20] For the moment, imagine

[18]Related to this is the empirical work of Lin and Png (1998), who find evidence to support the hypothesis that the importance of kinship in Chinese culture facilitates the use of contracting—as opposed to direct ownership—as a means to resolve opportunistic behavior in direct investments in the People's Republic of China.

[19]This normal-form "trust" game is similar in spirit to the extensive-form "trust" game that Kreps introduces in his Figure 1.

[20]As Kreps reminds us, this is equivalent to imagining the game is infinitely repeated with certainty but there is financial discounting.

that both the boss and the employee live at least as long as the game is repeated. Consider the following strategies for the players: The boss treats the employee fairly every period, unless she's exploited him in the past, in which case she continues to exploit him in this and every future period; and the employee trusts the boss, unless she's ever exploited him in the past, in which case he plays *don't trust* in this and every future period. Note, in the first period, these strategies call for the boss to treat fairly and for the employee to trust. Is this an equilibrium? Without going into all the details of the theory of repeated games (see, e.g., Gibbons, 1992, §2.3, for an introduction to this theory), the answer is *yes* provided the expected payoff to the boss of treating fairly outweighs the gain today from exploiting a trusting employee, but having to play against a untrusting employee forever after; that is, provided

$$\sum_{\tau=0}^{\infty} \beta^{\tau} \times 7 \geq 9 + \sum_{\tau=1}^{\infty} \beta^{\tau} \times 4.$$

A little algebra reveals that this condition holds if $\beta \geq \frac{2}{5}$ (i.e., if the probability of a game next period is at least 40%). Intuitively, provided she's never exploited the employee, the boss perceives the future expected benefit of being fair as outweighing the benefit of exploiting the employee given the strategy she anticipates the employee is playing. Note that, if it's credible the boss will treat fairly (i.e., if $\beta \geq \frac{2}{5}$), then the employee does better to trust than not to trust. Moreover, off the equilibrium path (i.e., if the boss has exploited), the employee does better not to trust given that he now anticipates the boss will always exploit him. In summary, we've seen that the fair treatment–trust outcome can be consistent with rational, self-interested, play by both parties in a repeated game in which $\beta \geq \frac{2}{5}$.

In this model, the employee need not be a long-lived player.[21] Since period by period he's simply playing his best response to what the boss is supposed to do given the history of the game, he could live only one period *provided each generation's employee knows the history of the game to that point.* The boss, on the other hand, must be long-lived. Although, similar to the junior-senior analysis of Figure 3 or the game considered by Kreps, she need not be infinitely lived—some overlapping generations story in which each generation's boss sells her stake in the firm to the next generation's boss is sufficient to make each generation's boss sufficiently concerned about the long run that cooperative play (i.e., the efficient outcome) can be sustained. For example, suppose at the end of each period, the boss can sell the firm to a new boss, the proceeds from which will fund the old boss's retirement. Suppose that each generation's employee expects his boss to treat him fairly if every boss in the firm's history has done so, but otherwise he expects to be exploited. The expected value, V_F, of a firm in which bosses treat their employees fairly is

$$V_F = \frac{7\beta}{1-\beta}.$$

[21] Again see Crémer (1986) for another example of a game among overlapping generations within an organization.

Suppose that there is a fairly competitive market in which to sell the firm, so that the old boss gets $V_F - \varepsilon$ for it; where $\varepsilon \geq 0$, but not too large. On the other hand, the expected value of a firm in which a boss has exploited an employee, V_E, is only

$$V_E = \frac{4\beta}{1-\beta}.$$

Again suppose a boss selling such a firm gets $V_E - \varepsilon$. Consider a firm in which bosses have so far treated their employees fairly. This generation's boss will continue this tradition provided today's payoff from doing so plus the greater value of the firm exceeds the payoff from exploiting plus the lower firm value; that is, provided

$$7 + V_F - \varepsilon \geq 9 + V_E - \varepsilon.$$

A little algebra reveals that this is equivalent to our original condition,

$$\sum_{\tau=0}^{\infty} \beta^\tau \times 7 \geq 9 + \sum_{\tau=1}^{\infty} \beta^\tau \times 4.$$

Hence, as advertised, we will still observe the efficient outcome even with a one-period-lived employee and an overlapping-generation boss (assuming $\beta \geq \frac{2}{5}$). This last point helps to make this a story about *corporate* culture: Observe that the reputation for treating employees fairly resides with the firm, not an individual. Corporate reputations or traditions can, thus, sustain desirable outcomes.

Observe that "fair treatment" and "exploitation" are fairly amorphous concepts. Typically, whether a behavior is "fair" is specific to the situation. For instance, giving a bad office to a poorly performing employee could be seen as fair, but giving the same office to a good performer could be seen as unfair. *Ex ante,* it is difficult to foresee all the possible behaviors that could be judged fair or unfair and, likewise, to enumerate all the situations in which they are fair or unfair. Here is where unforeseen contingencies arise. What is fair and what is exploitation cannot be defined in advance, yet, under the rules of a given culture, *ex post* behavior can be judged fair or not. In many ways, the situation is like pornography in the U.S.—material is legally obscene if it violates community standards and whether it, in fact, violates community standards is judged *ex post* under the criterion that people know obscene material when they see it.[22] Corporate culture thus enters into this model in two ways: One as reputation or tradition and two as the means of defining compliance with the reputation or tradition when future circumstances are difficult or impossible to define *ex ante*.

Of course, if "fair treatment" or "exploitation" cannot be defined in advance, but must be judged *ex post* according to the prevailing culture, one might also

[22]To paraphrase the "I know it when I see it" standard for pornography set down by the U.S. Supreme Court Justice Potter Stewart. Personally, the definition I've favored—which is not original to me—is pornographic material is material read with one hand.

imagine that the payoffs are equally ill-defined. That is, what the precise payoffs in the game shown in Figure 4 could be unforeseen.

Although the precise actions and payoffs could well be unforeseen, the *consequences* are not *necessarily* unforeseen. An example may help illustrate what I mean. A firm lures away a top manager from a rival. Before moving, he attempts to reach firm agreements on as much as he can with his new employer. But he knows that situations can arise, which he failed to anticipate, in which he would want his new employer to take an action that benefited him but was costly to his new employer. As a specific example, imagine he does not foresee that his current daycare arrangement will fall apart and that the best alternative daycare provider will require him to pick up his child half an hour earlier than before. This, in turn, means he must leave work half an hour early. Although he's willing to make up the work by arriving earlier in the morning or by working at home, such a schedule imposes a cost on the employer (e.g., because of difficulty in scheduling meetings). The employer can bear that cost or it can require he stay until 5, although this means a worse daycare arrangement for his child. Or the employer can use the employee's request as an excuse to hold up the employee and renegotiate aspects of the employment contract. Again we have the "trust" game of Figure 4: The employee can "trust"—go to a new firm knowing that he'll be vulnerable—or "not trust"—stay put; likewise the employer can accommodate the employee—let him adjust his schedule without penalty—or "exploit" him—force him to stay until five or otherwise give something up. Whereas the employee couldn't anticipate that he would want to change his schedule (i.e., specific actions and payoffs), he could anticipate that he would be relying on his new employer to treat him fairly (i.e., that he would face games like that in Figure 4). Part of his decision process, then, in deciding to switch employers was assessing the reputation (or culture if you will) of his potential new employer: How would he do in games like the one in Figure 4?

The fact that the consequences are foreseeable and, at a general level, the type of actions (fair treatment or exploitation) as well, raises the question of why contracts aren't used. Why can't the parties write a contract that says that the boss will treat the employee fairly? The obvious—but seriously incomplete—answer is that it would be difficult for a judge or other dispute adjudicator to determine whether treatment was fair or not. That is, a judge's assessment of how the employee was actually treated is subject to error. However, as Hermalin and Katz (1991, 1993a) and Edlin and Hermalin (1999) demonstrate, imprecise assessments do *not* preclude efficient contracting: As long as the judge's assessment is correlated with what the boss actually did, the parties can rely on this assessment as follows. They write a contract calling for the boss to treat the employee fairly and, after the boss has acted, they renegotiate the contract in anticipation of what might happen should they go to court. Under fairly general conditions (see, e.g., Edlin and Hermalin, 1999), the parties can construct a contract that duplicates the outcome they would have enjoyed were the boss's action verifiable without error. That is, as a matter of contract theory, there is no need to rely on culture at all, the parties can instead rely on formal contracts.

Despite having contributed to all three of these papers, I don't, in fact, believe that contracting is generally a substitute for culture.[23] First, as a matter of law, some of these contracts might not actually be enforceable. Second, more importantly, the costs associated with negotiating, monitoring, and enforcing these contracts could make them a less desirable means of inducing fair treatment—reputation can be equally effective and cost less. This is where the fact that inducing cooperation through repeated interaction is cheaper than inducing it contractually comes into play.

To summarize: Repeated play is a substitute for contracting.[24] Although not necessarily a perfect substitute, repeated play is typically a less costly way of inducing cooperative behavior (i.e., behavior that is not optimal in a one-shot situation). So far, culture hasn't entered in. Where culture enters in these models is to define the actions: Culture determines which actions are considered fair treatment and which are considered exploitation. More generally, culture defines what constitutes cooperative play—actions that are to be rewarded by future cooperation—and non-cooperative play—actions that are to be punished by withholding future cooperation. In this sense, culture substitutes for the impossible task of specifying all contingencies in advance. An employee may not, for instance, know in advance that cooperative play means allowing him to change his schedule to meet his childcare needs —but the culture will specify it as such should that particular contingency arise.

2.2 A Critical Assessment of Kreps

The previous discussion was intended to show that Kreps's theory of corporate culture is really two theories. One is a theory of culture as a way of ensuring coordination in games like those in Figures 1(a) and 2. The other is a theory of culture as a way of categorizing future contingencies for the purposes of sustaining cooperative play.

The first view could be called the culture-as-convention view. From a welfare perspective, it doesn't matter in the game of Figure 1(a) whether both players play A or they both play B. What is important is that they coordinate. In that sense, the game is analogous to whether we drive on the right (as in the U.S.) or on the left (as in the U.K.): It doesn't matter which we choose as long as we're coordinated.

A convention is, of course, only one way to ensure coordination. An alternative is to negotiate and reach some agreement.[25] For example, each morning

[23] Although the idea is not wholly insane either: There are many lawsuits by employees alleging unfair treatment; i.e., alleging that the boss violated the (implicit) employment contract.

[24] Schmidt (1995), citing his own work with Monika Schnitzer, points out that the *possibility* of formal contracting could destroy the possibility of using reputation when the two are close substitutes. Reputation is more sustainable the greater the difference between the cooperative and the non-cooperative payoffs. If non-cooperation results in switching to contracts and contracting is a close substitute for cooperation, then this difference is small and, consequently, reputation could be *un*sustainable.

[25] Given that an agreement is a Nash equilibrium (no one wants to unilaterally deviate from coordinating), an oral agreement is all that's necessary. For instance, if we all decided to drive

we could all discuss and agree on which side of the road we were to drive that day. However, as is immediate from this far-fetched example, in many contexts the costs of communicating to establish the coordinated outcome are large and a convention would, therefore, yield considerable savings (this anticipates some of the ideas in Crémer, 1993). Alternatively, the coordinated outcome could be dictated by some authority (e.g., a national driving committee). But except for one-off decisions, such as on which side to drive, this again involves costs: The authority must be apprised that coordination is needed, must assess the situation, and must transmit its decision. A general convention—such as defer to seniors—would save on these costs.

As we've just seen, a general convention saves considerably on the costs associated with explicit coordination. On the other hand, it, itself, is not without costs. By their nature, cultural conventions depend on people being aware of them, correctly knowing how to apply them, and being willing to enforce them. Hence, particularly when the corporate culture is unique, considerable expense could be incurred in socializing new employees (related to an earlier discussion, this may favor "buying" the prevailing national, regional, or professional culture over "making" a culture). In addition, reliance on convention over communication increases the odds of misunderstanding: For instance, two people playing the game of Figure 1(a) could each believe he or she was senior and—since they're relying on the convention, not communication—each chooses the action associated with the coordination outcome best for him or her (i.e., one plays A and one plays B). Kreps is silent on this issue, but clearly understanding these tradeoffs between the benefits and cost of culture is important. Future research should be devoted to both modeling these tradeoffs, as well as measuring them through field research.

Another issue left unexplored by Kreps is where these coordinating conventions come from. Recent work by Young (1993, 1998) has taken an adaptive learning approach to the issue: Players are not fully rational, but can adapt their play as a function of past experiences. As Young (1998, pp. 34–36) shows, if the proportion of past experiences of both players playing A in the Figure 1(a) game is high enough, then adaptive learning will converge to both players playing A in every period; i.e., a convention will have been established. Likewise, if the proportion of past experiences in which B was played was high enough, then a convention of B playing will be reached. On the other hand, it is possible that if the proportions of times A and B are played are just right, then players oscillate between A and B, there is no coordination, and a convention is not reached.[26] Of course, if the players find themselves oscillating—or more

on the right, then we don't really need a formal contract since no one would unilaterally want to deviate from the agreement (i.e., cause a head-on collision).

[26]Suppose each player treats the proportions of A and B played by his opponent as his opponent's mixed strategy and plays a best response to it. Suppose that Junior played B and Senior played A in the first round. In the next round, Junior would play A and Senior would play B. In the third round, Junior plays his best response to a mixed strategy of $(\frac{1}{2}, \frac{1}{2})$ and Senior does likewise. Hence, Junior plays B—which yields him an expected payoff of 1.5 versus an expected payoff of 1 were he to play A. Similar reasoning finds Senior playing A.

generally no convention emerges—then presumably the players will revert to communication and explicit coordinating activities to achieve coordination (or to set a convention if "A" and "B" can be sufficiently well defined in advance). In other words, to understand the use of a cultural convention *vis-à-vis* another solution may require some understanding of how conventions evolve or fail to evolve.

The second interpretation of culture offered by Kreps is that of culture as a way of categorizing future unforeseen contingencies for the purposes of sustaining cooperative play. Although Kreps's argument is clear, a proponent of formalism in economics might fault it on the grounds that there is too much "hand waving." In particular, there is no formalization of unforeseen contingencies. But Kreps is not to blame: There are no good formalizations of unforeseen contingencies in the economics literature (see Dekel et al., 1998, for a survey of the current state of modeling unforeseen contingencies).

Whether one can adequately formalize the culture-as-defining-cooperation interpretation is an open question. Certainly, partially adequate formalizations are possible, but even these require a certain amount of "hand waving" *vis-à-vis* the impact of unforeseen contingencies.[27] To the extent one is willing to put up with hand waving, progress is possible. To the extent one is unnerved by it, future economic research on this aspect of corporate culture is on hold until the profession develops acceptable models of unforeseen contingencies and, more generally, bounded rationality (see Rubinstein, 1998, for an overview of the current state of modeling bounded rationality in economics).

3 Other Economists on Corporate Culture

As noted in the introduction, Kreps is not the only economist to address the issue of corporate culture. In this section of the chapter, I consider the work of three other economists.

3.1 Crémer (1993)

Crémer (1993) ignores potential conflict among the actors within a firm. In his approach, all actors are perfectly honest and trustworthy. The tension in his approach arises from his assumption that these actors' capacity to process, receive, and transmit information is a scarce resource. Consequently, there is a payoff to economizing on communication and, therefore, a common stock of knowledge—what Crémer defines as culture—is valuable.[28] That is, as was

In the fourth period, Junior plays his best response to $(\frac{2}{3}, \frac{1}{3})$ and Senior to $(\frac{1}{3}, \frac{2}{3})$, which causes them to flip back. And the process continues indefinitely (provided we assume the correct tie-breaking rule should a player be indifferent between A and B; e.g., if we assume that Junior plays B when indifferent and Senior plays A when indifferent).

[27] An earlier version of this chapter considered a more explicit formalization of these issues, but it too required some hand waving.

[28] Crémer's precise definition of corporate culture is "the part of the stock of knowledge that is shared by a substantial portion of the employees of the firm, but not by the general population

implicit in Kreps (1990), culture substitutes for explicit communication. But Crémer's view of corporate culture is more nuanced than his explicit definition might suggest. In particular, he decomposes culture into three components (p. 362):

1. A common language or coding.

2. A shared knowledge of pertinent facts.

3. A shared knowledge of the norms of behavior.

The first two suggest insights not contained in Kreps.[29] Given the earlier discussion of the 3rd component in reviewing Kreps, here I will, consequently, focus on the first two components.

The importance of these two components can be illustrated by a recent discussion I had with a non-economist. She wondered if a growing demand for organic produce would raise the price of organic produce. I replied that since organic farming seemed likely to be a competitive industry, there would be a short-run price increase, but, in the long run, there wouldn't be a noticeable change in price. To an economist, this explanation is readily grasped.[30] But it took half an hour to explain my conclusions to this non-economist. As a non-economist, she didn't know the "pertinent facts" (e.g., that while the short-run supply curve in a competitive industry is typically upward sloping, the long-run supply curve, which is determined by the minimum of long-run average cost, tends to be flat; that is, the long-run price will equal the minimum of long-run average cost).[31] Related to this, she hadn't mastered the "common language and coding" (e.g., what a "competitive" industry is). The cost of our not having a common "economics" culture is the opportunity cost of the time I spent responding to her question.

We could have avoided this cost if I had helped her or encouraged her to acquire a knowledge of basic microeconomics. Given the infrequency with which she asks me economic questions, such an upfront investment would make little sense. Viewed in this light, to the extent that culture is an *ex ante* investment to lower later communication costs, it becomes something to which economics can speak directly.

---

from which they are drawn" (p. 354).

[29] Although Kreps's discussion of common ways of sorting American cities—see page 8 of this chapter—certainly relates to the shared knowledge of pertinent facts.

[30] Although an economist would understand my explanation, this doesn't, admittedly, mean he or she would agree with it. In particular, he or she might dispute whether organic farming is a competitive industry. In addition, he or she might point out that even if the farmers have no market power, there may be parties in the distribution channel (e.g., wholesalers or supermarkets) that do have market power, and, hence, a shift out in demand could lead to a price increase at the cash register.

[31] To be precise, the long-run supply curve is flat provided that entrants are as efficient as established firms. In the case of organic farming, this strikes me—admittedly on the basis of little information—as a reasonable approximation.

As an example, suppose that an employee needs to know in which rectangular region a given state occurs (e.g., in which sales area she should concentrate). Suppose there is a commonly understood coordinate system for identifying points in the relevant space (e.g., longitude and latitude or kilometers north and east of a prominent landmark). Then to identify a region, two pairs of coordinates must be sent, identifying diagonally opposed vertices of the rectangle in question; that is, a total of four numbers must be sent.[32] Suppose it costs c to send a number. Imagine there are four possible regions. An alternative to sending the four numbers that identify each region each time is to assign each of the four regions a single number ("name") and to first communicate to the employee the "code book" (i.e., region i is defined by coordinates (x_1^i, y_1^i) and (x_2^i, y_2^i)). This requires an initial investment of $20c$. But for each subsequent communication, there is a savings of $3c$ since, now, only one number (the region number) must be sent. Hence, provided the employer anticipates sending at least 7 messages, it pays her to invest in this "culture."[33] Fewer than seven messages, and a culture doesn't pay.

In real firms, of course, the employee isn't always given a cultural code book on day one. Rather, new employees are often given the information indirectly. That is, there is gradual cultural diffusion rather than immediate diffusion. This could be because indirect methods are cheaper. For example, there is "free" learning of pertinent knowledge and coding that occurs by listening to the conversation of experienced colleagues at lunch (although there is an opportunity cost to having an "uncultured" employee during this period of gradual diffusion). In many circumstances, it would actually be too difficult to identify *ex ante* all the pertinent knowledge and its coding. In addition, even if a "code book" could be written, it would often be the case that a new employee couldn't absorb it all at once.[34] Hence, there will often be no realistic alternative to gradual diffusion. But even if gradual diffusion makes sense, some rates of diffusion will make more sense than others. Viewed in this way, cultural diffusion becomes a dynamic programming problem, something with which economists are quite familiar. That is, a possible direction for future research in the economics of corporate culture would be to theorize about optimal cultural diffusion rates.[35]

[32]There are alternative means of communicating the region (e.g., sending the southwest coordinate plus the length & width), but these all require four numbers.

[33]To be precise, future savings should be discounted. Doing so would not, however, change the nature of the argument.

[34]Although strong-culture firms make a point of conveying as much of the "code book" explicitly in an initial period following an employee's hiring (see, e.g., O'Reilly and Chatman, 1996). Further, with respect to employees' ability to absorb it "all at once," this would depend on a number of factors, including their cognitive ability, existing values at entry, the amount of information, etc.

[35]The model of Garicano (1999) is somewhat related to this research plan. He, however, focuses on a static model that, in essence, considers how knowledge should be allocated among a firm's personnel to balance the cost of education against the cost of communication. In this sense, his approach is similar in spirit to the *formal* model in Crémer (this formal model does not, however, do justice to the richness of the ideas Crémer offers verbally).

Folk wisdom holds that the Inuit language has many different words for snow to allow the Inuit to make fine distinctions among different kinds of snow. The explanation is that snow is important to the Inuit. Similarly, in a firm where distinctions over certain states are important, we might expect a rich "language" to develop to help make these distinctions. As a newcomer to such a firm, I might rationally infer from this that I should pay particular attention to these states. That is, learning the language teaches me what I should consider important. But recognizing this, a firm might rationally want to develop a rich language—even when distinctions are not important—because it wants its employees to take things more seriously (note this relaxes Crémer's assumption of honest actors). Hence, a firm that really believes these distinctions are important will over invest in its language to signal that it really does hold these distinctions important.

To be somewhat more concrete on this signaling idea, imagine that there are two types of firms, each of which must deal with states in some space $\mathbf{S}$. Let $\mathcal{P}$ be a partition of $\mathbf{S}$ and let n be the number of elements (subsets of $\mathbf{S}$) in $\mathcal{P}$. A language can be seen as giving names to the elements in $\mathcal{P}$ and a richer language is one that corresponds to a finer partition (i.e., one with a higher value for n). Let $\mathcal{P}_B$ be the partition associated with the "base" language, which has a vocabulary of b words. Suppose this base language is sufficient for both types of firms. A firm can add w words to the language (employ a finer partition than $\mathcal{P}_B$) at a cost of cw. For one type of firm, distinguishing the states is not that important. In particular, an effort of e by a worker on distinguishing states returns only ℓe. For the other type of firm, distinguishing the states is important. For this type, an effort of e returns he, where $h > \ell$. Suppose to succeed in the first type of firm, a worker must expend at least e_ℓ effort, while he would need to expend at least e_h at the second type, where $e_h > e_\ell$. Workers dislike effort, but want to succeed. Consequently, if workers could distinguish the types of firms, they would expend e_h in the "h" type firm and e_ℓ in the "ℓ" type firm. Finally, allowing w to be a continuous variable,[36] there must exist a $\hat{w}$ such that

$$\ell e_h - c\hat{w} = \ell e_\ell.$$

What we should, then, expect in equilibrium is that an h-type firm will add $\hat{w}$ words to its language to signal that it is an h-type firm: Workers believe that a firm with fewer than $b + \hat{w}$ words is an ℓ-type firm and expend only e_ℓ in effort, but one with $b + \hat{w}$ or more is an h-type firm and expend e_h in effort. Given these beliefs, it is not rational for the ℓ-type firm to add $\hat{w}$ words, since the cost just cancels the benefit (fooling workers into working harder). But it is rational for the h-type firm, since the above equation implies

$$he_h - c\hat{w} > he_\ell$$

[36]The analysis when w must be a whole number is similar, but slightly more complicated. Hence, for convenience, I will treat it as continuous (a real number).

(recall $h > \ell$).[37] Observe that adding $\hat{w}$ words is overkill—by assumption b words were sufficient. This point could be generalized: To the extent that aspects of the culture cause new employees to make inferences about what is expected of them, we can easily imagine that some firms are induced to adopt an overly strong culture due to signaling considerations. To the best of my knowledge, this idea has not been explored in the literature on corporate culture.[38]

The reference to "code books" also suggests that culture can *worsen* agency problems. From novels and movies, we are all aware of secret signals and gestures used (in fiction, at least) by criminals to communicate without drawing the attention of the law-abiding or the police. There is no reason a similar phenomenon can't arise within a firm. Specifically, if a common culture reduces the costs of communication and, hence, coordination, then a *sub*-culture could reduce the costs of communication and coordination for a sub-group within the firm (e.g., lower-level employees) that wishes to conspire against another group (e.g., management). This suggests that the collusion-among-agents problem considered by Tirole (1992) (among others) could be exacerbated by culture (or sub-culture). This negative aspect of culture within corporations has not, as far as I know, been explored within economics.

In summary, it would seem that there is much that "off-the-shelf" economic theory can contribute to the shared knowledge and language aspects of corporate culture. It can, for instance, aid in studying these as investment problems (static or dynamic), the signaling aspects of culture, and the agency consequences of culture.

3.2 Hodgson (1996)

Hodgson rejects the methodology of both neo-classical economics (specifically contract theory) and the "new" institutional economics (e.g., Coase, 1937; Williamson, 1975). His primary criticism is that these approaches take the representative actor as unchanging in her preferences, attitudes, and modes of thinking as she moves from one situation to the next. Hodgson argues that a more appropriate assumption is that these aspects of her personality are shaped by her situations; that is, how she views a situation, thinks about it, and even what she wants to achieve in it evolve as she becomes immersed in the situation. It is not, as in standard economics, that only the actor is acting: The situation is, in essence, simultaneously acting on the actor. Note, too, that in Hodgson's view, the situation changes the actor not only by providing her information and skills—concepts that can be found in many neo-classical models—but also, most

[37] For more on signaling models in general, see, e.g., Gibbons (1992), §4.2.

[38] There is work in organizational behavior that has similarities to this, namely work on the symbolic nature of language (see, e.g., Pfeffer, 1981). This work points out that "managers in organizations send signals about what is important and valued through mechanisms such as what they spend time on, what they ask questions about, *and what they talk about* (Pfeffer, 1997, p. 125, emphasis added). This literature does not explain, in terms of rational actors, why these mechanisms are effective, nor the extent to which they would be deployed in equilibrium.

		Junior A	Junior B
Senior	A	$3, 2$	$0, -k$
	B	$0, 0$	$2, 3-k$

Figure 5: Coordination game between "cultured" junior and senior.

critically, by causing her preferences themselves to change.[39]

Two quotes from Hodgson illustrate the consequences of this view for understanding corporate culture:

> Corporate culture is more than shared information: through shared practices and habits of thought, it provides the method, context, values, and language of learning, and the evolution of group and individual competences (p. 255).
>
> [Institutions] play an essential role in providing a cognitive framework for interpreting sensedata (*sic*) and in providing intellectual habits or routines for transforming information into useful knowledge. A result of the framing or cognitive effects of institutions is to promote conformism, or emulation—to use Veblen's term. The availability of common cognitive tools, as well as perhaps a congenital or learned disposition for individuals to conform with other members of the same group, work together to mold and harmonize individual goals and preferences. Significant shifts in preferences and goals are involved, and such outcomes are an important part of the institutional self-reinforcing process (p. 263).

In Hodgson's opinion, corporate culture serves to mold the individual actor's preferences, attitudes, and modes of thinking.

To illustrate this view, recall the game of Figure 1(a). When a junior joins the firm, his preferences are as shown there. But as time passes, he "buys into" the prevailing corporate culture, with the consequence that he feels he ought to defer to seniors. For instance, not deferring causes him discomfort. Suppose that this discomfort costs the junior k units of psychological well-being, then culture has transformed the Figure 1(a) game into the game shown in Figure 5. If the discomfort from violating the deference norm is great enough—specifically, if $k > 3$—then A becomes a dominant strategy for the junior and the unique Nash equilibrium of the game is for both players to play A. Observe that this formulation of what culture does achieves the same outcome as Kreps's, but it does so in a different way. In Kreps's formulation, the junior defers to the senior not because that's what he inherently wants, but because he expects that the senior will play A and he does better to coordinate than not to coordinate. It's possible, in fact, that he even resents this outcome, since he knows that he would do better under the alternative B–B equilibrium. In contrast, Hodgson presumes that culture causes the junior to inherently prefer to play A. That is, he has "bought into" the norm to such a degree that he prefers to play A

[39] As a general rule, economists prefer to model preferences as fixed and exogenous. For exceptions, see Benhabib and Day (1981) or Hermalin and Isen (1999).

irrespective of strategic considerations. In this formulation, there is no question of the junior resenting the outcome; he has been programmed to like it.

Some observations are in order. First, if $k \leq 3$, then B–B remains an equilibrium. In other words, if the norm is not strong enough, then the same coordination problem considered in Section 2 re-emerges. In this case, it is not the change in the junior's preferences induced by the defer-to-seniors norm that achieves coordination, but rather the *mutual* expectation that juniors will defer under this norm. That is, even when norms change preferences, it could still be their impact on expectations that makes them valuable.

This raises the issue of how one could distinguish empirically between Hodgson's view and Kreps's. As we've seen, both could independently explain an observed defer-to-seniors norm. Moreover, even if an attitudinal survey revealed some internalization of the norm, that could still fail to be the explanation for the norm's effectiveness (e.g., when $k \leq 3$). Conversely, internalization could be what creates the expectation: Suppose that the population of juniors is split evenly into those who have strongly internalized the norm ($k > 3$) and those who have weakly internalized the norm ($k \leq 3$). A senior who meets a junior at random will play A—her expected payoff from playing A is, at worst, 1.5 versus, at best, 1 from playing B. But understanding this about seniors, even weak norm-internalizers will want to play A. Internalization has created expectation.

One might imagine that the two views could be better distinguished in situations resembling the game of Figure 4: If we have fair treatment and trust with*out* repetition between the players, then this would seem to suggest that the boss, at least, has internalized some fairness norm. Unfortunately, the controlled experiment that would test this is difficult to observe: By their very nature, interactions within firms tend to be repeated. Moreover, even if the boss doesn't play repeatedly with a given employee, the fact that she plays repeatedly with a sequence of employees, each of whom is likely to have some knowledge of her past history, can be sufficient for a reputation to be established (i.e., as noted previously, a single long-lived player can be sufficient for reputation to be effective; see §9.2 of Fudenberg and Tirole, 1991). Finally, although Hodgson is somewhat vague on how norms become internalized, one must imagine that it is through long-term exposure to the prevailing culture, which creates obvious problems for distinguishing between a game-theoretic view and a norm-internalization view.

This argument is not meant to cast doubt on norm internalization. Most of us tip at restaurants, are courteous to strangers, and don't take things that don't belong to us (even when detection is impossible). Although all these phenomena can explained in terms of game theory, in particular using Kandori's (1992) "contagion" model,[40] such explanations typically operate at the edge

[40]The basic idea in Kandori (1992) is that cooperative behavior (e.g., tipping, courteousness to strangers, not stealing, etc.) is an equilibrium in which people play cooperatively unless they encounter a non-cooperative person (e.g., a non-tipper, a rude person, discover they've been robbed, etc.), after which they too play non-cooperatively. Hence, even if I will never meet my current opponent again and have no fear that my non-cooperativeness will be directly revealed to others, I may wish to cooperate rather than triggering a collapse of the social norm: By serving as a non-cooperative "contagion," I start an epidemic of non-cooperation and,

(or over) of what is plausible game-theoretic reasoning. For some phenomena, Occam's Razor requires us to favor norm internalization over game-theoretic explanations.

Moreover, there is no reason to necessarily view norm internalization and game theory as antagonistic explanations. In many ways, they could be complementary. As noted earlier, internalization can generate game-theoretic expectation. Conversely, game-theoretic expectation might generate internalization: Constant exposure to the defer-to-seniors norm may ultimately lead a junior to believe he actually prefers to defer. That is, my casual empiricism suggests that humans often adapt to the situations in which they find themselves by convincing themselves that they are in a situation that they would have chosen.[41] This is, perhaps, because they wish to avoid believing they've made a bad choice or are poor decision makers (O'Reilly and Chatman, 1986).

This last discussion raises more fundamental questions. If norms are internalized, then the question becomes how? Why do people internalize norms? And what is the mechanism by which this occurs? Hodgson is relatively silent on these matters. The best he can offer is "a congenital or learned disposition for individuals to conform with other[s]." And even the second possibility, "a learned disposition," is not wholly satisfactory because one can ask why do individuals learn this? Why are they prone to learning conformism rather than non-conformism?

A genetic disposition to conformism seems a more promising hypothesis. There has been some work in economics arguing that certain human behaviorial traits make sense from an evolutionary perspective (e.g., Frank, 1988), but, with the exception of the growing field of evolutionary game theory (e.g., Weibull, 1995),[42] this has largely been outside the economics mainstream. Certainly, though, one can tell an evolutionary tale to make sense of a tendency to conform. Consider social animals such as fish that swim in schools or zebras that roam in tight herds on the African savannah. The advantage of being together (conforming) is that it protects the individuals from predators. Those individuals who were more prone to conform would, then, have a higher survival probability and, hence, more off-spring than those less likely to conform. That is, conformity would provide an evolutionary advantage. But this is not the end of the story: There needs to be some biological mechanism that steers these creatures towards conformity. Presumably, something akin to anxiety or fear strikes these creatures when they break away from the school or herd. Therefore, it was the tendency to anxiety from non-conformism that was the biological mechanism selected for.[43] Current humans and other hominoid species are so-

eventually, I'll start finding myself playing against non-cooperative players, to my detriment.

[41]The "Stockholm" phenomenon, whereby hostages begin to identify with their captors, could be seen as an extreme example of this tendency.

[42]Evolutionary game theory in economics, in turn, builds on earlier work by biologists (see, e.g., Maynard Smith, 1982; Hofbauer and Sigmund, 1988) that applied game-theoretic ideas to explain evolutionary forces.

[43]Weidensaul (1999) provides some indirect evidence for such a mechanism: Two differences between domesticated dogs and their wolf ancestors are (i) the former are less timid and (ii)

cial and it seems plausible, therefore, that these same biological mechanisms were selected for in the family *Hominoidea*. After all, breaking norms often induces a sense of unease or tension in people.[44] In short, like fish and zebras, we're programmed to follow the herd.[45]

Bernheim (1994) offers a model of conformity somewhat consistent with these ideas: People have inherent preference for status (perhaps because status once afforded a superior likelihood of reproductive success). People whose behavior indicates a propensity to obey norms achieve greater status than those whose behavior indicates a propensity to disobey norms. People, therefore, choose to conform. However, as in Kreps (1990), "conformity" in Bernheim's model has a strategic rather than intrinsic motivation—people don't gain intrinsic satisfaction from conforming, instead they dislike conforming *per se*, but prefer to be perceived as conformists rather than nonconformists. That is, unlike that for which Hodgson is arguing, norms are not internalized in Bernheim's model.

3.3 Lazear (1995)

Lazear (1995) can be seen as an attempt to formalize the process by which culture comes to be internalized. Lazear takes an evolutionary approach: Preferences are like a genetic endowment. At each moment, t, an individual in a firm meets ("mates") with another individual. This meeting causes each individual to produce an offspring: His or her $t+dt$ self, whose preferences ("genetic makeup") is a mixture of his or her former preferences and those of the individual he or she met (the "parents'" preferences). Through manipulation by top management, some preferences ("genetic endowments") are favored; that is, their carriers are more likely to survive to "mate." In essence, although without her precision, top management is like a horticulturist selecting for desired traits in flowers. Lazear's model could, therefore, be seen as one of *artificial* selection rather than *natural* selection.

For the most part, the close genetic analogy that Lazear pursues (down to each individual possessing two preference alleles, A and B—for three possible genetic endowments, AA, AB, and BB) is cute, but unnecessary. His ideas can be conveyed more straightforwardly—and arguably in a manner more consistent with actual behavior—as follows. Suppose that there are two possible prefer-

they produce different levels of the hormone thyroxine. These two differences are arguably related: Thyroxine affects the adrenal gland and the fear response is, in part, controlled by adrenaline. In the domestication process, creatures who were less timid around humans would be selected for; that is, creatures whose thyroxine levels produced lower adrenal responses would be selected for. As it turns out, many of the differences between dogs and wolves—in particular, dogs' pædomorphosis *vis-à-vis* wolves—are all directly or indirectly controlled by thyroxine.

[44]For more on the evolution of human behavior, see Goldsmith (1991).

[45]Within economics, there are also information-based models for herd behavior (see, e.g., Scharfstein and Stein, 1990; Banerjee, 1992; Bikhchandani et al., 1992). In essence, these models show that an individual will follow the herd because he draws a statistical inference that what the popular behavior is must also be the behavior that others have found is personally most rewarding.

Table 1: Proportion of Organization, $p(t)$, who are "As" (to five digits)

$t =$	$s = .1$	$s = .2$
1	.52498	.54938
5	.62246	.73106
10	.73106	.88080
15	.81757	.95257
50	.99331	.99995

ences (beliefs, mind-sets, etc.), A and B. Of these, A is the preference that top management wishes to promote. Assume the organization has many individuals, of which $p(t)$ proportion hold preference A and $1-p(t)$ hold preference B at time t. At each moment in time, individuals in the organization are randomly paired off (meet). Let $s \in [0, 1)$ denote top management's promotional effort in encouraging the adoption of A. Suppose this works to cause s proportion of Bs who meet As to come to hold preference A. Hence, of the $1 - p(t)$ of the organization's population that hold preference B, $p(t)$ of them will meet an A, and a further s of those will switch from B to A. Hence, at each instance in time, the proportion of As is increasing by $sp(t)[1 - p(t)]$; that is,

$$\dot{p}(t) = sp(t)[1 - p(t)].^{46}$$

Note that the greater is s, the faster the proportion of As increases. If we imagine that the organization begins at time 0 equally divided between As and Bs, then this differential equation has the solution:

$$p(t) = \frac{e^{st}}{1 + e^{st}},$$

where e is the base of the natural logarithm (i.e., $e \approx 2.7183$). Table 1 gives some values for $p(t)$. Suppose that the organization's cost today of choosing s is $C(s)$, an increasing and convex function. Suppose that the instantaneous benefit at time t to the organization from having proportion p prefer A is $V(p)$, where $V(\cdot)$ is increasing on the interval $(\frac{1}{2}, 1)$. If r is the interest rate, then the organization chooses s to maximize

$$\int_0^\infty V[p(t)]e^{-rt}dt - C(s).$$

For example, suppose that A and B refer to strategies in the game in Figure 6. Let $V(p)$ be the expected average payoff to members of the organization; that is,

$$V(p) = 3p^2 + 2(1 - p)^2 = 5p^2 - 4p + 2. \tag{1}$$

If $C(s) = 20s^2$ and $r = .05$ (i.e., the interest rate is 5%), then numerical calculations reveal that the optimal s is approximately .37. It can be readily

[46]Recall that a dot over a function indicates its time derivative; e.g., $\dot{p}(t) = dp(t)/dt$.

		Player 2	
		A	B
Player 1	A	$3, 3$	$0, 0$
	B	$0, 0$	$2, 2$

Figure 6: Another coordination game.

shown that if the interest rate, r, *falls* or the payoff in the A–A cell of the Figure 6 game increases, then this value for s will increase.

Note that in Lazear's formulation, as here, s is an upfront investment. One could also conceive of continual effort to promote the A-culture. In this case, the firm chooses a sequence of investments, $s(t)$, to maximize

$$\int_0^\infty (V[p(t)] - \hat{C}[s(t)])e^{-rt}dt$$

(where $\hat{C}(\cdot)$ is an increasing and convex function) subject to

$$\dot{p}(t) = s(t)p(t)[1 - p(t)].$$

Although the first-order conditions for this program are readily calculated, they are not particularly informative in and of themselves. Manipulating them, one can show that $\dot{s}(t) < 0$ *in a sufficiently old organization* (i.e., when t is large); that is, beyond some time we can be sure that the organization devotes continually less resources to promoting its culture. In contrast, for *young* organizations, it is not clear whether investments in culture promotion are steadily decreasing or increasing in the period following its birth.

Although we now have a formalization of the *consequences* of the internalization process, this is not the same as explaining the internalization process itself. Why do individuals switch from A to B? And what activities are represented by s? Lazear (p. 108) suggests some answers:

- Suppose that when an A meets a B there is some probability, σ_1, that the A complains to management about the B. Suppose, too, that there is some probability, σ_2, that management acts on that complaint and replaces the B worker with an A worker (i.e., fires the former and hires the latter). Here $s = \sigma_1 \times \sigma_2$. By rewarding "snitches" or otherwise encouraging low tolerance of Bs, management can work to raise σ_1. By increasing its efforts to replace identified Bs with new workers it identifies as As, management can work to raise σ_2.

- Suppose that workers play the game in Figure 6. When a payoff of 0 is realized (i.e., when a B meets an A), the B worker says to himself, "Something went wrong here." With probability s, he then recalls the training seminar, the distributed literature, or a motivational speech telling him that this is an A organization. Realizing his mistake, he then switches to playing A. By enhancing the effectiveness of the training seminar, the

salience of the literature, or the frequency of speeches, management can raise s.

Note that neither explanation really addresses internalization of preferences *à la* Hodgson (1996): Under the first explanation, no one's preferences change—workers with "bad" preferences are simply replaced with workers with "good" preferences. This is a Stalinist rather than Maoist approach (i.e., execution over re-education). Under the second explanation, all workers want to do the "right" thing—i.e., their interests coincide—but they only realize that they're "misbehaving" when things go terribly wrong (e.g., they get 0); at which point they re-assess their behavior. Observe that this explanation is more consistent with the convention aspect of culture than with the norm aspect of culture.

This discussion raises a couple of important issues. First, are norms internalized? Or are people holding certain norms selected for? These questions indicate that cross-sectional studies (e.g., comparing the attitudes of individuals at "strong" and "weak" culture organizations) are likely to shed little light on internalization. Rather, empirical work would have to be based on panel data (i.e., a longitudinal aspect is necessary), with care being taken to handle attrition in a statistically appropriate way.[47]

The second issue is what is being internalized. Or, in a similar vein, what does it mean to internalize a norm. As an American, have I internalized driving on the right or is driving on the right—given that I know the convention—merely a manifestation of my desire to protect myself, my passengers, and my car? If the former, then an approach like Hodgson or Lazear's is appropriate. If the latter, then it would seem preferable to stay with the more game-theoretic approach of Kreps (1990) or, possibly, Young (1993, 1998).

Even staying with the Lazear approach, observe that the two elements in Lazear's model, managerial effort to foster a culture, s, and "mating," are each, independently, sufficient to generate the type of dynamics that he considers. For instance, let σ be the probability that a B worker is "caught" and replaced with an A worker. Then

$$\dot{p}(t) = \sigma[1 - p(t)].$$

Consequently,

$$p(t) = 1 - Ke^{-\sigma t},$$

where K is a constant determined by the initial condition (e.g., if $p(0) = \frac{1}{2}$, then $K = \frac{1}{2}$). If the initial cost of setting up a monitoring system with σ effectiveness is $\tilde{C}(\sigma)$, where $\tilde{C}(\cdot)$ is increasing and convex, then the optimal monitoring effectiveness, σ^*, is the solution to

$$\max_{\sigma} \int_0^{\infty} V[p(t)]e^{-rt}dt - \tilde{C}(\sigma).$$

[47] See Chatman (1991) for an empirical study along these lines. She followed new employees of large accounting firms over a $2\frac{1}{2}$-year period. Her findings suggest that there is both socialization (internalizing) and selection at work.

Hence, if $V(\cdot)$ is given by equation (1), $\tilde{C}(s) = 20s^2$, $p(0) = \frac{1}{2}$, and $r = .05$, then $\sigma^* \approx .35$. A further examination of the dynamics shows that the value of $p(t)$ under this dynamic and the one in Lazear is exceedingly similar for $t \geq 5$ and σ or $s \geq .3$

Alternatively, suppose that as a holder of minority opinion or preference, a worker is inclined to change to the majority opinion or preference with probability $\phi(p)$, where $p > \frac{1}{2}$ is the proportion holding the majority opinion. Assume that if $p > \frac{1}{2}$, then $\phi(p) > 0$ and $\phi'(p) > 0$. This model attempts to capture the idea that people are prone to conform and that the pressure to conform is greater the more conformists with whom one comes in contact. Then

$$\dot{p}(t) = \phi[p(t)][1 - p(t)].$$

Suppose, e.g., that $\phi(p) = \alpha p$ for $p > \frac{1}{2}$ ($\alpha \in (0, 1)$), then

$$p(t) = \frac{e^{\alpha t}}{\kappa + e^{\alpha t}},$$

where κ is a constant determined by the initial condition; that is,

$$p(0) = \frac{1}{1 + \kappa}.$$

Hence, $\kappa = \frac{1}{p(0)} - 1$; thus,

$$p(t) = \frac{p(0)e^{\alpha t}}{1 - p(0) + p(0)e^{\alpha t}}.$$

For instance, if $p(0) = .6$ and $\alpha = .2$, then $p(10) \approx .92$. In this example, we have $dp(t)/dp(0) > 0$ and $dp(t)/d\alpha > 0$—the greater the initial proportion of the majority initially, the greater it is at any future point; and the more responsive the minority is to the majority, the greater is the majority at any future point. Note this model is similar to Lazear's, except here $p(0) > \frac{1}{2}$ and α is exogenous. Of course, if management can influence α, then we're back to Lazear's model (although the derivation is different).[48]

An alternative is that α (more generally, $\phi(p)$) is a function of the importance of having a single culture. Hence, for instance, when coordination is more important, then α could be larger. That is, the more important conforming is, the more rapidly we conform.

Although, as this section illustrates, there are many ways to formalize the dynamics by which culture might be propagated, we are still left with the problem of understanding the underlying process. Is it inter-personal contacts (i.e., "mating") that spreads culture? And, if so, what leads to one person's culture displacing another's? To what extent, if any, is management able to influence the spread of culture? Indeed, is management spreading culture or simply

[48]Note, too, that we can readily conceive of other functional forms for $\phi(\cdot)$. For instance, we might want $\lim_{p \downarrow \frac{1}{2}} \phi(p) = 0$, in which case $\phi(p) = \alpha(p - \frac{1}{2})$ would be a better model.

weeding out the non-believers? Lazear demonstrates that economics can help us understand the propagation and diffusion of culture within a firm;[49] but, as this discussion demonstrates, economics is unlikely to *explain* the underlying mechanics by which these processes operate—for these, economic modelers must rely on their sister social sciences or even the biological sciences.

4 Complementary Insights from the Economics of Organizations

The discussion so far has concerned economists' writing on corporate culture *per se*. In this section, I turn to the question of how results from the economics of internal and industrial organization could be used to enhance our understanding of corporate culture.

4.1 The Intersection between IO and Corporate Culture

To build on Lazear's (1995) analogy between culture and genes, the criterion for judging the desirability of a given culture is the corresponding fitness of the firm. That is, a culture can be judged only by the competitive advantages it yields the firm. In turn, to appreciate competitive advantages requires some understanding of the firm's competitive situation, which brings us to the area of microeconomics known as industrial organization. This section explores some ways in which industrial organization can complement an understanding of the importance of corporate culture and the diffusion of strong cultures within in an industry.

4.1.1 A Model of Universal Adoption or Non-Adoption

From earlier discussion in Sections 2 and 3, a reasonable assumption would seem to be that a strong culture leads to a more efficient organization; that is, one with lower *marginal* costs. But, as considered above, there is an overhead or fixed cost associated with instilling and maintaining a strong culture.[50] To formalize this, let a "cultured" firm's cost of producing x units be $F_c + m_c x^2$. Similarly, let the cost for an "uncultured" firm be $F_u + m_u x^2$. Consistent with

[49] Young (1993, 1998) and Bikhchandani et al. (1992), among others, also illustrate the ability of economics to model propagation and diffusion of aspects of culture such as conventions and fads.

[50] In this and subsequent discussion, I'm abstracting from the issue of *how* a firm instills and maintains a culture. I'm simply assuming that (i) it can (at a cost) and (ii) efforts to instill and maintain a culture are always successful. The second of these is, admittedly, unrealistic, but the conclusions reached in the analyses that follow would be relatively unaffected if we were to assume stochastic success. The *difficulty* of the analysis would, however, increase. In addition, by "strong culture" I mean culture that is both strategically appropriate and firmly accepted by the people within a firm. In this sense, I'm conflating the concepts of *cultural content* and *cultural strength*. See Flynn and Chatman (2000) in this Handbook for a further discussion of these concepts.

earlier discussions, assume that

$$0 < m_c < m_u \text{ and } 0 < F_u < F_c;$$

that is, the cultured firm has lower marginal costs, but higher fixed costs, than the uncultured firm.

Finally, suppose that this is a perfectly competitive industry. This means (see, e.g., Mas-Colell et al., 1995, §10.F) that in long-run equilibrium the only type of firm that survives is the type with the lower minimum average cost. Average cost is $mx + F/x$, which achieves a minimum at $x = \sqrt{F/m}$. Substituting this value of x into average cost yields

$$\text{AC}^{\min} = 2\sqrt{mF}.$$

If $\text{AC}_c^{\min} < \text{AC}_u^{\min}$, then only cultured firms will survive in long-run equilibrium. If the inequality is reversed, then only uncultured firms will survive. A little algebra reveals that the inequality—and hence an equilibrium with only cultured firms—will hold if

$$\frac{m_u - m_c}{m_u} > \frac{F_c - F_u}{F_c};$$

that is, if the proportional reduction in marginal costs from having a culture is greater than the proportional increase in fixed costs from having a culture. If the proportional reduction is less than the proportional increase, then only uncultured firms will exist in long-run equilibrium.

Some observations on this analysis. First, observe that

$$\sqrt{\frac{F_c}{m_c}} > \sqrt{\frac{F_u}{m_u}}.$$

This means that, in a cultured-firm equilibrium, each firm produces more than in an uncultured-firm equilibrium *ceteris paribus*.[51] Hence, in a cultured-firm equilibrium there will be fewer firms than in an uncultured-firm equilibrium *ceteris paribus*. Consequently, controlling for total industry production, we should observe fewer firms in a competitive industry in which corporate culture seems prevalent than in one in which it seems less prevalent. Since output correlates strongly with other measures of size, such as employees, it should also be that firms tend to have more employees when culture is a prevalent phenomenon than when it's not, all else equal.

With respect to empirical work, this prediction could cause problems when testing Lazear's (1995) prediction that a corporate culture is easier to instill in a small firm and that, therefore, smaller firms will tend to exhibit stronger corporate cultures. That the cost of instilling culture could be less the smaller is the firm indeed seems plausible (more on this momentarily), but the validity of Lazear's "therefore" also depends on how the *benefits* of culture vary with size

[51] Recall that, in the long-run equilibrium of a perfectly competitive industry with homogeneous firms, each firm operates at the minimum of its average cost curve (see, e.g., Mas-Colell et al., 1995, §10.F).

(a point he does not address). If, as here, benefits are assumed to increase with size (since the benefit of reducing marginal costs is greater the more that will be produced), then one must compare whether the increase in benefits outweighs the increase in costs.

To aid in understanding this point, suppose, as a change to what was previously assumed, that the cost of x is

$$F + \frac{x^2}{s},$$

where s, an endogenous variable, is the expenditure per worker on instilling and maintaining culture. As before, expenditures on culture reduce the marginal cost of production. Suppose that the production technology is such that each worker produces γ units. Hence, if L is the number of workers, then γL units are produced and the firm's total cost is

$$F + sL + \frac{\gamma^2 L^2}{s}$$

when expressed in terms of its workforce (size). Note that the marginal cost of culture increases with size, L. But the marginal benefit does as well:

$$\frac{d^2}{dL\,ds}\left(-\frac{\gamma^2 L^2}{s}\right) = \frac{2\gamma^2 L}{s^2} > 0.$$

For a given L, the firm chooses s to minimize total cost; hence, the optimal s is $\gamma\sqrt{L}$. This is an *increasing* function of L, meaning that larger firms will invest more *per worker* in culture than smaller firms.

This isn't, however, a fully general conclusion. Suppose the cost of instilling and maintaining culture in a firm with L employees were sL^α. Then the firm would choose s to minimize

$$sL^\alpha + \frac{\gamma^2 L^2}{s}.$$

The solution is

$$s = \gamma L^{\frac{2-\alpha}{2}}.$$

This function is increasing in L for $\alpha < 2$—as in the previous paragraph, larger firms invest more per worker in culture than smaller firms. If, however, $\alpha > 2$, then this conclusion is reversed: Smaller firms would then invest more per worker. Finally, if $\alpha = 2$, then investment per worker is independent of size. This analysis, thus, validates the earlier point that how culture varies with size depends critically on how the benefits and costs of culture vary with size. And while the *costs* of culture can be largely studied by looking at the firm in isolation, the *benefits* of culture will typically require some understanding of the firm's competitive environment.

		Z: Adopt	Z: Don't
Y	Adopt	$\frac{(a-m_c)^2}{9b} - I, \frac{(a-m_c)^2}{9b} - I$	$\frac{(a+m_u-2m_c)^2}{9b} - I, \frac{(a+m_c-2m_u)^2}{9b}$
	Don't	$\frac{(a+m_c-2m_u)^2}{9b}, \frac{(a+m_u-2m_c)^2}{9b} - I$	$\frac{(a-m_u)^2}{9b}, \frac{(a-m_u)^2}{9b}$

Figure 7: Culture Adoption Game between Firms Y and Z .

4.1.2 An Asymmetric Model of Culture Adoption

The competitive model lends itself to equilibria with homogeneous behavior.[52] On the other hand, observation suggests that many industries are characterized by heterogeneous behavior. If we presume some link between culture and behavior, then such observations suggest that we should observe *intra*-industry heterogeneity in culture. Industrial organization can help explain this heterogeneity.

Building on earlier work of mine (Hermalin, 1994), consider an industry with just two firms, Y and Z. The timing of the model is that, first, the firms decide whether to invest in instilling and maintaining a corporate culture or not. Let I be the cost of such an investment if undertaken. As before, the benefit of a corporate culture is that it lowers marginal cost. Let m_u be the marginal cost of a firm that lacks a strong culture and m_c be the marginal cost of a firm that has a strong culture, where $m_c < m_u$. After establishing a culture or not, each firm observes the strength of its rival's corporate culture. Then the firms decide how much to produce. Let industry demand be such that $p = a - bX$, where p is the market price, X is total industry output, and a and b are fixed, positive parameters. Assume competition between these two firms is Cournot competition (see, e.g., Mas-Colell et al., 1995, §12.C).[53] Consequently, if Y's marginal cost is m_Y and Z's is m_Z, then the firms' profits (gross of investments in culture) are

$$\pi_Y = \frac{(a + m_Z - 2m_Y)^2}{9b} \text{ and } \pi_Z = \frac{(a + m_Y - 2m_Z)^2}{9b},$$

respectively. Consequently, the culture-adoption game is the one shown in Figure 7. Observe that

[52]To be precise, if $m_cF_c = m_uF_u$, then a heterogeneous equilibrium would be possible. Alternatively, if $m_cF_c < m_uF_u$ but there was an exogenous *limit* on the number of cultured firms, then the marginal firm could be uncultured and the equilibrium would be heterogeneous. These possibilities likely arise too infrequently for them to serve as adequate general explanations of heterogeneous equilibria.

[53]Under Cournot competition, firms simultaneously choose their output. The price they receive is $p = a - bX$. Note that each firm's output imposes a negative externality on the other firm by reducing the price that the second firm will receive.

$$(a+\tilde{m}-2m_c)^2-(a+\tilde{m}-2m_u)^2=\int_{m_c}^{m_u}4(a+\tilde{m}-2m)\,dm.$$

Consequently, since the integrand is increasing in $\tilde{m}$, it follows that

$$\frac{(a-m_c)^2}{9b}-\frac{(a+m_c-2m_u)^2}{9b}<\frac{(a+m_u-2m_c)^2}{9b}-\frac{(a-m_u)^2}{9b}. \qquad (2)$$

In words, the lower is your rival's marginal cost, the smaller the gain from reducing your own marginal cost. Intuitively, the lower your rival's marginal cost, the greater your rival's output; and the greater your rival's output, the smaller will be your output in the Cournot equilibrium, which means marginal cost reduction is less valuable to you.

A consequence of expression (2) is that there exists an interval of Is such that

$$\begin{aligned}\frac{(a-m_c)^2}{9b}-I &< \frac{(a+m_c-2m_u)^2}{9b} \text{ and}\\ \frac{(a+m_u-2m_c)^2}{9b}-I &> \frac{(a-m_u)^2}{9b}.\end{aligned}$$

For an I in that interval, the game in Figure 7 has only two pure-strategy Nash equilibria: In one, firm Y adopts or invests in a corporate culture, while firm Z does not; and, in the other, firm Z adopts or invests, while firm Y does not. Observe that we have heterogeneity in both equilibria: One firm has a strong culture, while the other does not.[54,55]

If the cost of having a culture is very low (lies below the interval), then there is a single Nash equilibrium in which both firms adopt a strong culture. Conversely, if the cost of having a culture is very high (lies above the interval), then there is a single Nash equilibrium in which neither firm adopts a strong culture.

Another comparative static is with respect to a: As demand shifts out (a increases), then the value of establishing a strong culture increases. Consequently, we can expect three regions: For low enough a, neither firm has a strong culture; for intermediate values, the equilibrium is heterogeneous; finally, for high values, both firms have a strong culture. This, in turn, suggests some empirical tests of this theory: For instance it should be that the prevalence of strong cultures is greater in industries with higher output per firm than in industries with less output per firm. That is, all else equal, we should see stronger cultures in larger firms.[56]

[54]For an I in this interval, there is also a third Nash equilibrium in mixed strategies. Since this equilibrium will also yield heterogeneity with positive probability, the prediction of heterogeneity can be said to hold for all of the game's equilibria.

[55]Some readers may worry that these heterogeneous equilibria are the consequence of assuming that culture is a discrete decision. By analogy with the analysis in Hermalin (1994), it can, however, be shown that these conclusions can also be reached in a model in which culture is chosen from a continuum of possible values.

[56]This conclusion is driven by the assumptions of the model: Marginal benefit of culture

With further regard to empirical work, this model suggests a puzzle for determining the long-run impact of culture. This model indicates that the strength of a firm's culture is a function of the size it expects to attain in equilibrium. That is, somewhat loosely, size determines culture. This, in turn, creates some issues for interpreting correlations between culture and firm success: Could a positive correlation be spurious? Many measures of success, such as profits and longevity, are positively correlated with size. Could, then, the competitive opportunities that lead to large size explain both the strength of culture and the apparent success of the firm?

It is worth noting that one could also generate heterogeneous equilibria through other models. For instance, it is sometimes suggested that a strong culture makes the firm more efficient in a specific environment, but less efficient in other environments, relative to a firm with less culture. If there is sufficient uncertainty over future environments, then, staying with Cournot competitors, one can find equilibria in which one firm will adopt a strong culture, while the other will not (even if the cost of establishing a culture is zero).[57]

Another model that could generate heterogeneity would be a product-differentiation model. Suppose that instead of affecting marginal cost, culture raised the quality of the product or service of a firm (e.g., customers value dealing with a "service-with-a-smile" culture more). Consider a duopoly. If both firms adopt a strong culture or if neither does, then what the firms produce or provide will be perceived by consumers as homogeneous. In turn, this means that consumers will decide from which to buy solely on the basis of price, which leads to a form of competition known as Bertrand competition (see, e.g., Mas-Colell et al., 1995, §12.C). In Bertrand competition, economic profits are driven to zero by the ferocity of the price competition. If, instead, the firms differentiate their product or service along some non-price line (e.g., culture), then price competition is less fierce and the firms can make positive profits.

4.1.3 Lessons

This section has concerned itself with the importance of tying industrial organization into a study of corporate culture. To the extent that corporate culture is a choice variable for the firm, the level or intensity of a firm's culture depends on both the costs and benefits of a strong corporate culture. Whereas the costs can be reasonably understood by looking at the firm only, the benefits depend significantly on the firm's competitive environment. This, in turn, requires an examination of that environment and an understanding of the firm's strategic responses to it.

The nature of the conclusions that one can reach about corporate culture depend critically on the firms' competitive environment. Does competition en-

is increasing in size, while marginal cost is not (the cost of culture is a fixed cost). This conclusion would be reversed if the model were such that the marginal cost of culture was increasing faster with size than the marginal benefit.

[57] It is a reasonably well-known fact that Cournot competitors would prefer that their marginal costs be uncorrelated. The one cite, however, I know for this result is Dana (1991).

courage more or less culture? How might the strength of culture vary with the size of the firm? How do we interpret correlations between strength of culture and other firm attributes, such as performance? Answering these questions requires linking culture to industrial organization.

4.2 Fairness and Related Models

A reasonable characterization of most economic models is that the rational actor is motivated solely by her payoffs. In particular, she will like equally all allocation processes that yield the same payoffs to her and what these processes provide others is irrelevant to her. Over the years, a minority of economists have recognized the extremity of that assumption and analyzed the consequences of individual actors caring about the process by which they receive their payoffs or caring about what payoffs others get (consider, e.g., Veblen, 1899; Varian, 1974; Rabin, 1993; Frank, 1998). The extent to which these non-material concerns matter is both a manifestation of culture and an avenue through which manipulation of culture can affect behavior.

To appreciate how this might affect the study of culture, consider the following game between two individuals, Y and Z. Suppose that Y has \$100 and she decides to send ϕ dollars of it to Z, where $0 \leq \phi \leq 100$. Suppose that if she sends ϕ dollars, Z actually receives 3ϕ dollars. Z then decides how much, ψ, to send back to Y. The payoff to Y is $100 - \phi + \psi$ dollars and the payoff to Z is $3\phi - \psi$ dollars. Although this game is quite artificial, it reflects many situations in which one individual (here, Y) can gain if she trusts another (here, Z) who controls the returns from her investment (here, the process that turns ϕ dollars into 3ϕ dollars) to adequately share the proceeds of her investment with her.

If this game is played once and the players are fully self-interested, then Z would never return any money to Y. Since Y is rational, she would anticipate this and not send any money in the first place. Hence, in equilibrium, no money is sent. This is clearly worse for both players than a feasible outcome in which Y sends the entire \$100 and Z returns more than \$100.

In contrast, suppose that the players' utility functions are

$$\begin{aligned} U_Y &= \ln(w_Y) - \mu f(w_Y, w_Z) \\ U_Z &= \ln(w_Z) - \mu f(w_Z, w_Y), \end{aligned}$$

where, w_t is the final wealth of player t, μ is a positive constant, and, following Rabin, $f(\cdot)$ is a "fairness" function. Specifically, suppose that

$$f(w_1, w_2) = \begin{cases} \frac{w_1 - w_2}{100}, & \text{if } w_1 > w_2 \\ 0 & \text{, if } w_1 \leq w_2 \end{cases}.$$

That is, a player suffers some remorse if his final wealth exceeds the other player's because of the "unfairness" of the allocation.

Consider Z's play. If $\phi \leq 25$, then his final wealth can never exceed Y's, he can't feel remorse, and so he'll want to keep all of the 3ϕ he receives. In this

case, Y chooses her final wealth to maximize

$$\ln(w_Y) - \frac{\mu}{100}(w_Y - w_Z) = \ln(w_Y) - \frac{\mu}{100}(4w_Y - 300).$$

The solution to this program is $w_Y = \min\{25/\mu, 100\}$, provided $\mu \leq \frac{1}{3}$ (this proviso is necessary because we're considering the case where $w_Y \geq 75$). If, instead, $\phi > 25$, then it will be Z who can't have the smaller final wealth. Z will choose ψ to maximize

$$\ln(3\phi - \psi) - \frac{\mu}{100}(3\phi - \psi - [\psi + 100 - \phi]) = \ln(3\phi - \psi) - \frac{\mu}{100}(4\phi - 2\psi - 100).$$

The solution to this program is

$$\psi = \begin{cases} \max\left\{3\phi - \frac{50}{\mu}, 0\right\}, & \text{if } \frac{100}{\mu} \geq 2\phi + 100 \\ 2\phi - 50 & \text{, if otherwise} \end{cases}.$$

Observe, first, that if Y sends more than \$25, then Z will return some of it provided μ is large enough. Assume this is the case, then Y chooses ϕ to maximize

$$\ln\left(\max\left\{100 + 2\phi - \frac{50}{\mu}, 2\phi - 50\right\}\right).$$

The solution is $\phi = 100$, Y sends all her money. Putting it all together:

- If $\mu \geq \frac{1}{3}$, Y sends \$100 and Z returns \$150. That is, surplus is maximized when $\mu \geq \frac{1}{3}$.
- Define $\mu^* \approx .22696$.[58] If $\mu^* < \mu < \frac{1}{3}$, then Y sends \$100 and receives $300 - \frac{50}{\mu}$ dollars in return. Note she receives back more than \$100 if $\mu > \frac{1}{4}$ and receives less than \$100 back if $\mu < \frac{1}{4}$. Despite the inequity in the sharing, surplus is still maximized in the region $\mu^* < \mu < \frac{1}{3}$.
- If $\mu \leq \mu^*$, then Y sends nothing. Surplus is *not* maximized in this region.

From an organizational perspective, a firm would like, therefore, to ensure that $\mu > \mu^*$. That is, it would like to instill a culture that makes its employees sufficiently sensitive to fairness that Y can "trust" Z. There are many practices that might fit this bill. Examples are activities that build camaraderie, such as company picnics, an inviting staff lounge, company sports teams in local recreation leagues, and so forth. Alternatively, the organization could try to screen for fair-minded individuals or sanction employees who behave unfairly.

Observe that the kind of cooperation that occurs when $\mu \geq \frac{1}{4}$ could also be achieved through a repeated game of the sort considered by Kreps (1990).

[58] μ^* is the solution to the equation

$$\ln\left(300 - \frac{50}{\mu}\right) = \ln(100) - \mu.$$

Even if $\mu = 0$, an equilibrium can arise in a repeated game in which Y sends a \$100 each period and Z returns an amount greater than \$100. This equilibrium is supported by Y's threat to discontinue sending money should Z ever fail to return a sufficient amount to Y. In this case, we would be hard pressed to tell whether culture was sustained by repeated interaction among wholly selfish players or the internalization of a fairness norm. We could tell (at least reject wholly selfish players), however, if $\mu^* < \mu < \frac{1}{4}$—in this case, we would observe Y sending more than she received back, behavior that would be impossible if the players were wholly selfish.[59]

In a related vein, Kandel and Lazear (1992) consider the question of norm enforcement. These authors distinguish between *guilt*, which is internal, and *shame*, which requires observation by others. For instance, if I tip when traveling by myself far from home, then my tipping could be driven by guilt—I would feel bad if I didn't tip. In contrast, if I tip only when dining with others or only at restaurants I frequent often, then my tipping is probably driven by shame—I worry about the disapproval I would suffer from others.

The basic ideas in Kandel and Lazear's article can be captured using the following teams model (based on Holmström, 1982): Two employees (team members), Y and Z, each choose a level of effort, a_Y and a_Z respectively, that stochastically determines the team's output. Specifically, output, X, is $\zeta \times (a_Y + a_Z)$, where $\zeta = 1$ with probability q and equals 0 with probability $1 - q$ $(0 < q < 1)$. Suppose that any realized output is shared equally by the team members. The standard assumption in this type of modeling (see, e.g., Holmström, 1982) is that each member's utility has a form similar to

$$\mathcal{U} = \frac{1}{2}X - \frac{1}{2}a^2.$$

That is, it's equal to some function of his share of the output minus some function of his effort that represents his disutility of effort. Expected utility maximization would, then, lead each employee to choose $a = \frac{1}{2}q$. This, however, doesn't maximize *social* surplus: Each unit of effort returns only $\frac{1}{2}q$ in expected output to a worker; that is, he perceives his *private* marginal benefit of effort to be $\frac{1}{2}q$. However, his effort also benefits his team member. Adding that benefit in, we see that the *social* marginal benefit is q. Since this is larger than the private marginal benefit, we can conclude that each worker expends too little effort *vis-à-vis* the social optimum. Or, as economist might put it, neither worker values the positive externality that his effort has on his co-worker and, hence, expends too little effort.

In contrast to this standard formulation, suppose, first, that employees feel guilty if they expend less than first-best effort (i.e., if $a < q$), since they feel guilty about the harm their laziness causes their co-workers. Specifically, suppose that a worker's utility is

$$\widehat{\mathcal{U}} = \frac{1}{2}X - g(a) - \frac{1}{2}a^2,$$

[59] See Rabin (1993, 1998) for surveys of experimental evidence that supports rejecting the wholly selfish model of players in games like this.

where $g(\cdot)$ is the guilt function. For concreteness, suppose

$$g(a) = \begin{cases} G \times (q - a), \text{ if } a < q \\ 0 \qquad\qquad\quad , \text{ if } a \geq q \end{cases}$$

($G > 0$). The equilibrium expenditure of effort is, then,

$$a = \begin{cases} \frac{1}{2}q + G, \text{ if } G < \frac{1}{2}q \\ q \qquad\quad , \text{ if } G \geq \frac{1}{2}q \end{cases} . \tag{3}$$

Observe that guilt leads to a greater expenditure of effort than when workers aren't prone to feeling guilty. Moreover, the greater is G, up to $\frac{1}{2}q$, the more effort workers expend. As with fairness, there is, thus, an incentive for the firm to pursue activities that will make workers feel guilty if they slack off. In general, I imagine these activities would be similar to those that instill fairness (see above).

In contrast to guilt, let's now consider shame. Shame requires that a slacking worker's co-worker be able to detect that he has slacked. A co-worker has this ability only when $\zeta = 1$, since otherwise there is no output. When $\zeta = 1$, workers can determine each other's effort by subtracting their own effort from X. Suppose that a worker's shame is $g(a)$, where $g(\cdot)$ is the function defined in the previous paragraph. Then each worker's expected utility is

$$\widetilde{\mathcal{U}} = \frac{1}{2}X - qg(a) - \frac{1}{2}a^2.$$

In equilibrium, each worker's expenditure of effort is

$$a = \begin{cases} \frac{1}{2}q + qG, \text{ if } G < \frac{1}{2} \\ q \qquad\quad , \text{ if } G \geq \frac{1}{2} \end{cases} .$$

Comparing this to the equilibrium effort given by (3), we see that for any G, $G < \frac{1}{2}$, the workers expend less effort when motivated by shame than when motivated by guilt. This accords with Kandel and Lazear's more general finding that guilt is a stronger motivator than shame. But even if workers are motivated by shame, there is still a benefit to the corporation of trying to raise the amount of shame, G, they feel when caught. If, as seems reasonable, the amount of shame one feels is increasing in the number of people aware of his misconduct, then one way to raise G is by increasing the notoriety of those caught misbehaving.[60]

As with fairness, the behaviors attributable to guilt and shame can also be explained using a repeated-game framework. Each worker expends q in effort each period unless, in a previous period, aggregate output was something other than 0 or $2q$. In that case, he expends just $\frac{1}{2}q$ in effort. A high-effort equilibrium

[60]This could also be done indirectly by heavily publicizing the names of those who *behaved*, since co-workers can infer from this who didn't behave.

exists provided[61]

$$\begin{aligned}\frac{1}{2}q^2\frac{1}{1-\beta} &\geq \sum_{\tau=0}^{\infty}\beta^\tau(1-q)^\tau\left(\frac{5}{8}q^2+\frac{\beta}{1-\beta}\frac{3}{8}q^2\right)\\ &= \frac{q^2}{1-(1-q)\beta}\left(\frac{5}{8}+\frac{\beta}{1-\beta}\frac{3}{8}\right),\end{aligned}$$

where β is again the probability of continuing the game (alternatively, the financial discount factor). Provided $q > \frac{1}{2}$ and

$$\beta \geq \frac{1}{2(2q-1)},$$

repeated play will induce first-best effort from both workers even though they are not susceptible to guilt nor to shame.

Although these models demonstrate that fairness, guilt, and shame can be incorporated into economic analysis in a way that connects to corporate culture, there are a number of questions still left unanswered. Why do feelings about fairness, guilt, and shame affect human behavior? Why are, assuming they are, these feelings susceptible to manipulation (e.g., how does a company picnic make employees treat others more fairly)?

As might be expected, there has been little work in economics examining the origin of feelings such as fairness, guilt, or shame. To the extent these have found their way into economic models, it has typically been simply to assume they exist and, then, as above, consider their consequences. An exception to this is Frank (1987, 1988), which considers the evolutionary advantages provided by being subject to certain feelings, such as guilt. In particular, under appropriate assumptions,[62] a population with a conscience (being subject to guilt) can resist a small invasion of guilt-free "mutants"; moreover, a population that lacks a conscience is subject to being overrun by a small invasion of guilt-prone "mutants." That is, in the language of evolutionary game theory, an equilibrium in which people "choose" to have a conscience is evolutionary stable (see Maynard Smith, 1982; Weibull, 1995, for more on evolutionary stability). Yet showing that "guilt" could be evolutionarily advantageous is not, ultimately, to explain *why* we are subject to guilt. At some point, the underlying biological mechanism by which guilt works and how this mechanism evolved must be identified and understood.

On the other hand, social science is not a sub-discipline of biology (despite what some biologists might think). As long as we have evidence that people are subject to guilt (or shame or a preference for fair outcomes), then nothing prevents us from incorporating this into our models. Similarly, if we have evidence that these feelings can manipulated in ways suggested above, then we

[61] This expression takes into account that if a worker deviates, he does best to choose $a = \frac{1}{2}q$. Of course, his deviation might not be detected immediately because ζ could equal zero that period. This, too, is accounted for in the expression.

[62] See Harrington (1989) and Frank (1989) for a debate over just how reasonable these assumptions are.

are again free to incorporate such manipulation into our models of corporate culture. But this not a license for loose modeling—before we can take seriously a model that posits company picnics lead to employees treating each other more fairly, we need either evidence to support this or a better understanding of the bio-evolutionary mechanisms behind it.

4.3 Informal Authority: Leading & Delegation

Much of the economic modeling of organizations considers situations of formal authority. One party, often dubbed the *principal*, has some formal authority to coerce another party, often dubbed the *agent*. This authority could represent the right to order the agent about or it could be the ability to set the agent's incentives.[63] Recently, however, the economics literature has begun to explore issues of *in*formal authority. Here, I will discuss two strands of the literature and their links to corporate culture.

4.3.1 Leadership

Leadership is an important topic in sociological and psychological studies of organization. It has, however, received considerably less attention from economists.

One notable exception is Rotemberg and Saloner (1993), which considers the consequences of the leader (boss) having empathy with those under her.[64] In many ways, the analysis in Rotemberg and Saloner is similar to the analysis of fairness and guilt considered in the previous section. For instance, in keeping with the spirit of Rotemberg and Saloner, the leader could be Z in the money-transfer game considered above. If the leader is empathic, then the underling, Y, will know that the leader will behave fairly, which means the underling can trust the leader to return a sufficient amount of money to him. Or, to tie this more closely to real firm behavior, the underling could make investments that payoff for him only if the leader later recognizes them by increasing the underling's wage.

Some recent work of mine (Hermalin, 1998, 1999) is another exception. Unlike Rotemberg and Saloner, I stay more within the mainstream of neo-classical economics by positing that all actors are motivated solely by their own interests. I begin with the observation that a person is a leader only if she has followers. Following is an inherently *voluntary* activity. Hence, the essential question becomes how does a leader induce others to follow her.[65] As an economist, I presume that followers follow because it is in their interest to do so. What could make it in their interest to follow? One answer is that they believe the leader has better information about what they should do than they have. Leadership

[63] Typically, this authority is not absolute: The agent is almost always assumed to have the right to quit (or at least not accept employment initially). Often he is assumed to enjoy limited-liability protection, thereby preventing the principal from fining him.

[64] Also consider Rotemberg and Saloner (1998).

[65] In this sense, I can be seen as modeling what Max Weber (Gerth and Mills, 1946) refers to as "charismatic" leadership.

is thus, in part, about transmitting information to followers. But this can't be all: A leader must also convince followers that she is transmitting the *correct* information; that is, she must convince her followers that she is not misleading them.

The formal model I consider is related to the teams model considered in the previous section. Loosely, imagine the same model (with wholly selfish workers), except, now, the two values ζ can take are $\frac{1}{2}$ and 1. We can also imagine the team has $N \geq 2$ members, in which case each team member gets $1/N$ of the output.[66] Observe that the greater a worker thinks ζ will be, the more effort he will expend. Now suppose one team member, the leader, learns what ζ will be. This knowledge is her private information. If she thought that the other workers would believe her announcement, she would have an obvious incentive to always claim that $\zeta = 1$, since then she induces the most effort from her fellow team members and she, recall, gets ζ/N of that additional effort from each fellow team member. Of course, her fellow team members aren't naïve, they understand this temptation, so they rationally disregard her claims. But this is inefficient, since effort should be conditioned on its true marginal return, ζ. To overcome this, the leader must convincingly signal her information. Hermalin (1998) considers two methods of doing so: One, the leader can "sacrifice"; that is, give an *ex ante* gift to her followers when it is the good state (i.e., $\zeta = 1$). Two, the leader can "lead by example." In leading by example, the leader chooses her effort first and publicly. Based on her effort choice, the followers are able to infer what ζ must be and condition their efforts accordingly.[67]

Both Rotemberg and Saloner's work and my own can be related to corporate culture. In the case of Rotemberg and Saloner, the obvious relationship is culture as a mechanism to induce or enhance empathy. In my work, there are a number of avenues to explore. First, suppose that ζ reflects the importance of adhering to a cultural norm and effort now means a worker's effort to abide by the norm. Leading by example, then, reflects the oft-given advice that a leader should "walk the talk." That is, followers infer the importance of adhering to the norm by observing the degree to which the leader adheres. Somewhat along the same lines, workers—particularly those new to the organization—could be seeking to identify the organization's norms. They may naturally look to a single person, the leader, as a model of appropriate behavior. Here, again, it's critical that the leader "walk the talk." Conversely, it is often claimed that part of being a leader is inducing others to break with the past. Now the cost of following the leader could be the disutility caused by violating old norms.

[66] In Hermalin (1998), I show that, dividing output equally is the optimal arrangement in many contexts (but not all). Even when it's not, the arguments presented here continue to hold when the team uses the most efficient division.

[67] Hermalin (1998) considers only a one-shot model. Hermalin (1999) extends the analysis to a repeated game, where I show that leader sacrifice and leading by example could still be important. In a repeated setting, two more mechanisms for inducing the leader to be honest can emerge: (1) the leader can simply develop a reputation for honesty; and (2) the leader can be induced to be honest by the promise of tribute from her followers (see Hermalin, 1999, for details).

Followers' willingness to do so could be influenced by their inference of how much importance the leader places on this change. Again, leading by example or sacrifice could be critical.

Another aspect to leadership that I've considered, but not written on, would be to consider leadership in the context of labeling theory (Erikson, 1966). When explicitly defining norms is difficult, perhaps because they are complex or finely shaded, followers could infer the appropriate norms from observing who the leader rewards, and how much, and who she sanctions, and how severely. Given that followers are making such inferences, the leader could wish to reward or sanction in a strategic way; that is, with an eye towards influencing both what her followers infer and the speed with which they infer it.

4.3.2 Delegation

Delegation of authority is a common feature in all large organizations. Although various economists have studied aspects of delegation for years (arguably going back at least to Berle and Means, 1932), I will here discuss only Baker et al. (1999).

Baker et al. essentially consider a repeated version of the authority model set forth in Aghion and Tirole (1997). Their key insight is that authority can rarely be fully relinquished: The person who has the right to grant authority to another typically retains the right to rescind that grant at her choosing. If this person exercises that right opportunistically, then the advantages of delegation could be eliminated.

Formally, imagine that a principal hires an agent. By expending effort, a, the agent discovers, with probability qa, a project that could be undertaken ($0 < q < 1$). Assume the effort choice is binary: Either the agent expends effort, $a = 1$, or he doesn't, $a = 0$. With probability $\frac{1}{2}$, the project yields the agent a private benefit of B_b^A and with probability $\frac{1}{2}$, the agent's private benefit is B_g^A. Independently, the project yields the principal a benefit of B_b^P with probability $\frac{1}{2}$ and a benefit of B_g^P with probability $\frac{1}{2}$. Independence means that the four possible combinations of benefits are all equally likely. Assume that if there is no project, the benefit to each party is 0. The "b" benefit is bad and the "g" benefit is good in the sense that

$$B_b < 0 < B_g.$$

Initially only the agent knows if he's discovered a project and what the benefits are. Before, however, he can undertake a project, he must reveal the benefits to the principal. And although the principal has, in some sense, delegated to the agent the decision to proceed or not with a project, she could, at this point, rescind that authority and veto a project. Assume that the principal cares only about her benefit and the agent cares about benefit minus cost of effort, $C(a)$, where $C(1) = \bar{c}$ and $C(0) = 0$. To close the model, assume

$$B_b^P + B_g^A > 0$$

and

$$\frac{1}{2}qB_g^A > \bar{c} > \frac{1}{4}qB_g^A.$$

The first of these means that it is surplus maximizing to undertake a project that has a bad benefit for the principal, but a good benefit for the agent; in other words, it is surplus maximizing to undertake any project that's good for the agent. The meaning of the second expression will become clear in a moment.

Observe that the agent will suppress any project that would give him B_b^A. The expected social value, then, of the agent expending effort, *assuming the principal acts to maximize social surplus,* is

$$\frac{1}{2}q\left(B_g^A + \frac{1}{2}B_g^P + \frac{1}{2}B_b^P\right) - \bar{c}$$

(recall any project that's good for the agent should be pursued). By assumption, this quantity is positive; that is, the organization wants the agent to expend effort. However, in a one-shot game, there's no reason to expect that the principal will act to maximize *social* surplus rather than her own. This means she would veto (rescind authority) whenever the project would yield her B_b^P. Since, however, a project that's good for both parties occurs with only probability $\frac{1}{4}$, the agent would *not* expend effort if he anticipated this behavior by the principal—recall

$$\bar{c} > \frac{1}{4}qB_g^A.$$

Hence, the threat of opportunistic behavior by the principal results in a socially undesirable equilibrium in which the agent doesn't expend effort.[68]

If, however, the delegation game is repeated, then an equilibrium can exist in which the agent does expend effort each period: Provided the principal has never rescinded authority (vetoed a project), the agent expends effort. If the principal previously rescinded authority, the agent expends no effort. For an appropriately large value of β, this can be shown to be an equilibrium.

How does this relate to corporate culture? Note, as pointed out in footnote 68, one can both conceive of contractual solutions and convincingly argue against the assumptions necessary to rule out such solutions. The only possible merit to Baker et al.'s repeated-game solution is it's cheaper than using contracts. In essence, this returns us to the discussion in Section 2.1.2 (pages 12–13). That is, delegation could viewed as a specific application of the more general ideas in Kreps (1990). In particular, delegation could be supported by a culture of worker autonomy or voice. Actions by superiors that are perceived to violate such autonomy are violations of the culture.

[68]The reader may question this result, noting that there could be contractual solutions. That's right, but, following Baker et al. (1999), I'm assuming such solutions aren't feasible (benefits are observable, but not verifiable). Of course, this assumption could be criticized on the grounds raised by me and my co-authors in Hermalin and Katz (1991, 1993a) and Edlin and Hermalin (1999); but more on this later.

4.3.3 Lessons

I've intended the discussion in this section to illustrate that some recent work in the economics of organization may have a natural link with the study of corporate culture. I've also sketched out some ways in which this link could be explored or how an appreciation for corporate culture could aid in the interpretation of this work. It is, of course, beyond the scope of this chapter to consider the *entire* economics of organization, but if we did, I'm confident that we'd find many more potential links to and uses for notions of corporate culture.

5 Conclusions

This chapter has sought to examine how economics has been, to date, integrated into the study of corporate culture and how it might be, in the future.

For the most part, I've followed a game-theoretic approach to corporate culture. Following Kreps, I've focused on two kinds of games: Coordination games and repeated games in which cooperation is not an equilibrium of the one-shot (stage) game. The focus on coordination games leads to an emphasis on the convention-setting aspect of corporate culture. Having and understanding conventions can prevent coordination failures and economize on other means of ensuring coordination. Repetition can sustain cooperation in games in which cooperation would otherwise be lacking. In many such games, contracting could easily substitute for repetition. As I argued above, culture, as a means of defining appropriate behavior, is what could make repetition the better option: Explicitly defining appropriate behavior, particularly in a world with unforeseen contingencies, is difficult and costly. An implicit, culturally given, set of definitions would, therefore, be economizing relative to formal contracts.

The analysis of both these kinds of games leads to a view of culture as ultimately being a substitute for explicit communication. That is, culture is an unspoken language giving directives to the members of an organization (a view echoed by Crémer, 1993). Under this view, it might be better to write that a member of the organization *understands* the culture, rather than he or she is *part of* the culture. In particular, one's behavior represents the rational acceptance, based on the preferences one brings to the organization, of the cultural directives. This stands in contrast to the view, argued for by Hodgson (1996), Kandel and Lazear (1992), and Rabin (1993), that culture operates by changing one's preferences. In their view, culture is internalized.

Introspection suggests that much of culture must be internalized. My notions of female beauty are probably very similar to most other American males. Moreover, these notions are often at odds with the standard notions of different cultures (e.g., beauty as suggested by Rubens). Certainly among the thoughts I have when I see a beautiful woman is *not*, "Rationally, should I choose to obey the cultural directive to respond to this woman's beauty in the following ways . . . "

On the other hand, once we as economists begin fooling with preferences,

we risk losing the rigor imposed by neo-classical modeling conventions. An old joke illustrates what's at issue:[69]

> A philosophy professor wishes to validate Aristotle's claim that people are driven by the desire to be happy. "Mr. Smith," he asks, "what do you want?"
>
> "To make a lot of money."
>
> "Why?"
>
> "To buy things."
>
> "Why?"
>
> "Because that makes me happy."
>
> "Yes, Mr. Smith is driven by a desire to be happy," exclaims the professor triumphantly. "Ms. Jones, what do you want?"
>
> "To be a doctor."
>
> "Why?"
>
> "To help people."
>
> "Why?"
>
> "Because I enjoy helping people."
>
> "Yes, Ms. Jones is driven by a desire to be happy! And what about you Mr. Brown?"
>
> "I want to be sad," replies Mr. Brown with a sly look in his eyes.

If we were allowed to take Mr. Brown's claim at face value, then most of the conclusions of neo-classical economics would be turned on their heads. It is by sticking to a narrow and consistent set of assumptions that we are able to tease out interesting, plausible, and testable conclusions. Or, as John Freeman (1999) writes,

> All theory oversimplifies reality. The question is not so much what is left out, but how much can be explained with the simplest account. Adding variables or complications to the functional form imposes a cost on the theorist. That cost is the difficulty of falsification. If one throws everything that seems to matter into a theory, accounting for *every* observation and *every* anecdote, then falsifiability is threatened. (p. 174, emphasis added.)

Playing with preferences is, thus, a potentially dangerous activity. Like other things that are dangerous, but have important uses—fire comes to mind—caution is in order. In particular, to allow culture to shift individuals from being solely self-interested to partially group-interested is a risk. If not careful, we'll simply assume our conclusions or have a model that's so flexible that it can never be falsified. This is not to suggest that Kandel and Lazear (1992) or Rabin (1993) have been reckless, just that one following a similar approach must (i) exercise the same caution that these authors did and (ii) have strong evidence from experiments, other social sciences, or even biology, to support their assumptions.

[69] A version of this joke was originally told to me by the sociologist Marion Levy.

Ultimately, it is not economists' comparative advantage to try to resolve the "directive vs. internalization" issue.[70] First, there may be no resolution: Like light, which can be seen as both a wave and a particle, culture may best be understood from multiple perspectives. Second, were a resolution possible, it likely requires an understanding of psychology, particularly evolutionary psychology, that is beyond the standard training of most economists. What economists can do is to explore the consequences of these assumptions (e.g., as Kreps, 1990; Kandel and Lazear, 1992; Crémer, 1993; Rabin, 1993, do). Alternatively, they can seek to model the diffusion of culture as Lazear (1995) does or, for instance, by extending leadership models (e.g., Rotemberg and Saloner, 1993; Hermalin, 1998) as I suggested above.[71] Finally, they can employ their insights on costs and benefits, particularly those benefits derived from marketplace interactions, to investigate how organizations might optimally influence their cultures.

The economics of corporate culture is far from settled. Like the western U.S. after Lewis and Clark's expedition, this territory has only begun to be explored—to say nothing of developed.[72] I hope others will find this a fertile territory in which to work.

References

Aghion, Philippe and Jean Tirole, "Formal and Real Authority in Organizations," *Journal of Political Economy*, February 1997, *105* (1), 1–29.

Arrow, Kenneth J., *The Limits of Organization*, New York: W.W. Norton & Co., 1974.

Axelrod, Robert, *The Evolution of Cooperation*, New York: Basic Books, 1984.

Baker, George, Robert S. Gibbons, and Kevin J. Murphy, "Informal Authority in Organizations," *Journal of Law, Economics, & Organization*, March 1999, *15* (1), 56–73.

Banerjee, Abhijit V., "A Simple Model of Herd Behavior," *The Quarterly Journal of Economics*, August 1992, *108* (3), 797–818.

Benhabib, Jess and Richard H. Day, "Rational Choice and Erratic Behaviour," *Review of Economic Studies*, July 1981, *48* (3), 459–471.

Berle, Adolph A. and Gardiner C. Means, *The Modern Corporation and Private Property*, New York: MacMillan, 1932.

[70] For an additional economist's take on the issue of internalization, see Kreps (1997).

[71] As noted earlier, the work of Young (1993, 1998) and Bikhchandani et al. (1992) represent yet further means for economists to study cultural diffusion.

[72] Given that there were indigenous peoples living in the West at the time of Lewis and Clark's expedition, this, admittedly, might not be the most "PC" simile. On the other hand, the notion of an invasion might have considerable resonance with other, non-economist, social scientists.

Bernheim, B. Douglas, "A Theory of Conformity," *Journal of Political Economy*, October 1994, *102* (5), 841–877.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, "A Theory of Fads, Fashion, Custom, and Cultural Changes as Informational Cascades," *Journal of Political Economy*, October 1992, *100* (5), 992–1026.

Chatman, Jennifer A., "Matching People and Organizations: Selection and Socialization in Public Accounting Firms," *Administrative Science Quarterly*, September 1991, *36* (3), 459–484.

_ **and Karen A. Jehn**, "Assessing the Relationship between Industry Characteristics and Organizational Culture: How Different Can You Be?," *The Academy of Management Journal*, June 1994, *37* (3), 522–553.

Coase, Ronald, "The Nature of the Firm," *Economica*, 1937, *4*, 386–405.

Crémer, Jacques, "Cooperation in Ongoing Organizations," *Quarterly Journal of Economics*, February 1986, *101* (1), 33–50.

_ , "Corporate Culture and Shared Knowledge," *Industrial and Corporate Change*, 1993, *2* (3), 351–386.

Dana, James, "Differentiation and Emulation in Games of Strategic Choice under Uncertainty," 1991. Working Paper, Department of Economics, Dartmouth College.

Dekel, Eddie, Barton L. Lipman, and Aldo Rustichini, "Recent Developments in Modeling Unforeseen Contingencies," *European Economic Review*, May 1998, *42* (3–5), 523–542.

Edlin, Aaron S. and Benjamin E. Hermalin, "Renegotiation of Agency Contracts that are Contingent on Verifiable Signals," 1999. Working paper, Walter A. Haas School of Business, University of California.

Erikson, Kai T., *Wayward Puritans*, New York: John Wiley & Sons, 1966.

Flynn, Francis J. and Jennifer A. Chatman, "Strong Cultures and Innovation: Oxymoron or Opportunity," in Cary L. Cooper, Susan Cartwright, and Christopher Earley, eds., *Handbook of Organizational Culture (and Climate)*, Chichester, England: John Wiley & Sons, Ltd., 2000. Forthcoming.

Frank, Robert H., "If *Homo Economicus* Could Choose His Own Utility Function, Would He Want One with a Conscience?," *American Economic Review*, September 1987, *77* (4), 593–604.

_ , *Passions within Reason*, New York: W. W. Norton & Co., 1988.

_ , "If *Homo Economicus* Could Choose His Own Utility Function, Would He Want One with a Conscience?: Reply," *American Economic Review*, June 1989, *79* (3), 594–596.

_ , *Luxury Fever*, New York: The Free Press, 1998.

Freeman, John, "Efficiency and Rationality in Organizations," *Administrative Science Quarterly*, March 1999, *44* (1), 163–175.

Fudenberg, Drew and Jean Tirole, *Game Theory*, Cambridge, MA: MIT Press, 1991.

Garicano, Luis, "A Theory of Knowledge-Based Hierarchies: Communication, Organization and Technological Choice," 1999. Working paper, Graduate School of Business, University of Chicago.

Gerth, H. H. and C. Wright Mills, *From Max Weber: Essays in Sociology*, New York: Oxford University Press, 1946.

Gibbons, Robert S., *Game Theory for Applied Economists*, Princeton, NJ: Princeton University Press, 1992.

_ , "Game Theory and Garbage Cans: An Introduction to the Economics of Internal Organization," in Jennifer J. Halpern and Robert N. Stern, eds., *Debating Rationality: Nonrational Elements of Organizational Decision Making*, Ithaca, NY: ILR Press, 1998.

_ , "Taking Coase Seriously," *Administrative Science Quarterly*, March 1999, *44* (1), 145–157.

Goffman, Erving, *Stigma: Notes on the Management of Spoiled Identity*, Englewood Cliffs, NJ: Prentice-Hall, 1963.

Goldsmith, Timothy H., *The Biological Roots of Human Nature*, Oxford: Oxford University Press, 1991.

Harrington, Joseph E., "If *Homo Economicus* Could Choose His Own Utility Function, Would He Want One with a Conscience?: Comment," *American Economic Review*, June 1989, *79* (3), 588–593.

Hart, Oliver D., *Firms, Contracts, and Financial Structure*, Oxford: Oxford University Press, 1995.

_ **and Bengt Holmström**, "The Theory of Contracts," in Truman F. Bewley, ed., *Advances in Economic Theory Fifth World Congress*, Cambridge, England: Cambridge University Press, 1987.

Hermalin, Benjamin E., "Heterogeneity in Organizational Form: Why Otherwise Identical Firms Choose Different Incentives for their Managers," *RAND Journal of Economics*, Winter 1994, *25* (4), 518–537.

_ , "Toward an Economic Theory of Leadership: Leading by Example," *American Economic Review*, December 1998, *88* (5), 1188–1206.

_ , "Leading for the Long-Term," 1999. Working Paper, Walter A. Haas School of Business, University of California, Berkeley.

_ **and Alice M. Isen**, "The Effect of Affect on Economic and Strategic Decision Making," 1999. Working Paper, Walter A. Haas School of Business, University of California.

_ **and Michael L. Katz**, "Moral Hazard and Verifiability: The Effects of Renegotiation in Agency," *Econometrica*, 1991, *59*, 1735–1753.

_ **and** _ , "Defense Procurement with Unverifiable Performance," in Jim Leitzel and Jean Tirole, eds., *Incentives in Procurement Contracting*, Boulder, CO: Westview Press, 1993.

_ **and** _ , "Judicial Modification of Contracts between Sophisticated Parties: A More Complete View of Incomplete Contracts and Their Breach," *The Journal of Law, Economics, & Organization*, Fall 1993, *9* (2), 230–255.

Hodgson, Geoffrey M., "Corporate Culture and the Nature of the Firm," in John Groenewegen, ed., *Transaction Cost Economics and Beyond*, Boston: Kluwer Academic Press, 1996.

Hofbauer, Josef and Karl Sigmund, *The Theory of Evolution and Dynamical Systems*, Cambridge, England: Cambridge University Press, 1988.

Hofstede, Geert, Bram Neuijen, Denise Daval Ohayv, and Geert Sanders, "Measuring Organizational Cultures: A Qualitative and Quantitative Study across Twenty Cases," *Administrative Science Quarterly*, June 1990, *35* (2), 286–316.

Holmström, Bengt, "Moral Hazard in Teams," *Bell Journal of Economics*, Autumn 1982, *13* (2), 324–340.

Kandel, Eugene and Edward P. Lazear, "Peer Pressure and Partnerships," *Journal of Political Economy*, August 1992, *100* (4), 801–817.

Kandori, Michihiro, "Social Norms and Community Enforcement," *Review of Economic Studies*, January 1992, *59(1)* (198), 63–80.

Kreps, David M., "Corporate Culture and Economic Theory," in J. E. Alt and K. A. Shepsle, eds., *Perspectives on Positive Political Economy*, Cambridge, England: Cambridge University Press, 1990.

_ , "Intrinsic Motivation and Extrinsic Incentives," *American Economic Review*, May 1997, *87* (2), 359–364.

Lazear, Edward P., "Corporate Culture and the Diffusion of Values," in Horst Siebert, ed., *Trends in Business Organization*, Tübingen, Germany: J.C.B. Mohr (Paul Siebeck), 1995.

Lin, Chu-Chia S. and Ivan Png, "Kinship, Control, and Incentives," 1998. Working paper, Department of Economics, National Chengchi University, Taiwan.

MacLeod, W. Bentley, "Incentives in Organizations: An Overview of Some of the Evidence and Theory," in Horst Siebert, ed., *Trends in Business Organization*, Tübingen, Germany: J.C.B. Mohr (Paul Siebeck), 1995.

Mas-Colell, Andreu, Michael D. Whinston, and Jerry R. Green, *Microeconomic Theory*, Oxford: Oxford University Press, 1995.

Maynard Smith, John, *Evolution and the Theory of Games*, Cambridge, England: Cambridge University Press, 1982.

Morita, Hodaka, "Choice of Technology and Labor Market Consequences: Explaining U.S.-Japanese Differences in Management Style," 1998. Working paper, Department of Economics, Cornell University.

Okuno-Fujiwara, Masahiro, "The Economic System of Contemporary Japan: Its Structure and Possibility of Change," *Japanese Economic Studies*, 1994, *22*, 76–98.

O'Reilly, Charles and Jennifer A. Chatman, "Organizational Commitment and Psychological Attachment: The Effects of Compliance, Identification, and Internalization on Prosocial Behavior," *Journal of Applied Psychology*, 1986, *71* (3), 492–499.

_ and _ , "Culture as Social Control: Corporations, Cults, and Commitment," in Barry M. Staw and L. L. Cummings, eds., *Research in Organizational Behavior*, Vol. 18, Greenwich, CT: JAI Press, Inc., 1996.

Pfeffer, Jeffrey, "Management as Symbolic Action: The Creation and Maintenance of Organizational Paradigms," in L. L. Cummings and Barry M. Staw, eds., *Research in Organizational Behavior*, Vol. 3, Greenwich, CT: JAI Press, Inc., 1981.

_ , *New Directions for Organization Theory*, Oxford: Oxford University Press, 1997.

Rabin, Matthew, "Incorporating Fairness into Game Theory and Economics," *American Economic Review*, December 1993, *83* (5), 1281–1302.

_ , "Psychology and Economics," *Journal of Economic Literature*, March 1998, *36*, 11–46.

Rotemberg, Julio J. and Garth Saloner, "Leadership Style and Incentives," *Management Science*, November 1993, *39* (11), 1299–1318.

_ and _ , "Visionaries, Managers, and Strategic Direction," 1998. Working Paper, Stanford Graduate School of Business.

Rubinstein, Ariel, *Modeling Bounded Rationality*, Cambridge, MA: MIT Press, 1998.

Scharfstein, David S. and Jeremy C. Stein, "Herd Behavior and Investment," *The American Economic Review*, June 1990, *80* (3), 465–479.

Schmidt, Klaus M., "Comment on W. Bentley MacLeod, 'Incentives in Organizations: An Overview of Some of the Evidence and Theory'," in Horst Siebert, ed., *Trends in Business Organization*, Tübingen, Germany: J.C.B. Mohr (Paul Siebeck), 1995.

Tirole, Jean, "Collusion and the Theory of Organizations," in Jean-Jacques Laffont, ed., *Advances in Economic Theory: 6th World Congress*, Cambridge, England: Cambridge University Press, 1992.

Vandello, Joseph A. and Dov Cohen, "Patterns of Individualism and Collectivism across the United States," *Journal of Personality and Social Psychology*, August 1999, *77* (2), 279–292.

Varian, Hal R., "Equity, Envy, and Efficiency," *Journal of Economic Theory*, 1974, *9*, 63–91.

Veblen, Thorstein, *The Theory of the Leisure Class*, New York: MacMillan, 1899.

Weibull, Jörgen W., *Evolutionary Game Theory*, Cambridge, MA: MIT Press, 1995.

Weidensaul, Scott, "Tracking America's First Dogs," *Smithsonian*, March 1999, *29* (12), 44–57.

Williamson, Oliver, *Markets and Hierarchies: Analysis and Antitrust Implications*, New York: The Free Press, 1975.

Young, H. Peyton, "The Evolution of Conventions," *Econometrica*, January 1993, *61* (1), 57–84.

_ , *Individual Strategy and Social Structure*, Princeton, NJ: Princeton University Press, 1998.