

# **Toward a Theory of Corporate Culture: An Evolutionary Approach to Reciprocity in the Employment Relationship**

**Jessica L. Cohen\***  
The Brookings Institution

April 23, 2001

Version 0.5

**Preliminary and Incomplete**

\* I am extremely grateful to William Dickens for his extensive and insightful comments and suggestions (particularly his comments on the model). I would also like to thank Gil Skillman for assistance with this research as it appeared in my honors thesis. Generous support from the Surdna Foundation and the Davenport Committee at Wesleyan University is gratefully acknowledged. All errors are my own.

## Introduction

Since Akerlof's paper (1982) analyzing the role of reciprocity, or gift-exchange, in the firm, the case for the potential significance of this norm for the employment relationship has been bolstered significantly. A reciprocal employment relationship is one in which the worker offers a level of effort above the minimum-enforceable level, and receives above-competitive compensation in exchange. Economic experiments have consistently revealed that people are motivated by reciprocity in economic exchanges.<sup>1</sup> Survey evidence has documented the fundamental role that reciprocity plays in worker motivation and compensation.<sup>2</sup> Further, a good deal of this evidence of reciprocal behavior cannot be explained within the standard framework of strategically-rational decision-making. Although there have been attempts to analyze reciprocity as a strategic choice of behavior, most analyses of reciprocity and other cultural traits treat such motivations for behavior as given. Economic analyses typically conceive of a norm as a behavioral characteristic that is adopted by an individual through social learning that takes place outside of the economic sphere and is "carried with them" into their economic interactions. The development of the social norm is exogenous to these economic interactions, and behavior is analyzed by taking these motivations as given.

This conception of the source of reciprocal motivation leaves us with the following puzzle: If the adoption of a social norm is exogenous to economic interactions, why do we see reciprocal relationships systematically developing in certain types of economic exchanges and not others? How can it be, for instance, that gift-exchange employment relationships tend to develop only in certain types of firms? The notion that economic institutions may influence reciprocal behavior has received some attention in the past—for example, in economic experiments which have shown that the level of reciprocity changes when the experiment is explained to subjects in the context of an economic interaction (Hoffman, et al. 1994). Akerlof (1982) proposes that gift-exchange relationships will typically develop in primary sector firms, arguing that the rents to reciprocal employment relationships in the primary sector may be the cause of queuing and involuntary unemployment. Furthermore, there is an extensive literature in organizational psychology and sociology suggesting that employee motivation may be very sensitive to certain aspects of the work environment, such as the scope for employee participation, and the form of supervision and compensation.<sup>3</sup> Economic evidence of systematic productivity differences across certain types of firms, viewed in light of the theoretical arguments and evidence of reciprocity from the other social sciences, suggests that the organization of the firm may play a significant role in influencing the reciprocal behavior of its employees.

In the model presented here, the adoption of the social norm reciprocity is endogenized, and we analyze how the development of the reciprocal employment relationship may be sensitive to certain structural characteristics of firms. In an evolutionary game theory model, we illustrate how the adoption of the reciprocity norm

---

<sup>1</sup> Fehr and Gächter (2000), Güth (1995), Davis and Holt (1993) and Thaler (1988).

<sup>2</sup> Bewley (2000) and Kahneman, et al. (1986).

<sup>3</sup> See Guzzo and Katzell (1987), Miller and Monge (1986), Locke and Schweiger (1979) and Herzberg (1968) for reviews of the literature.

can be shaped by the costs and benefits to reciprocity that an individual perceives both within the firm and within society. Thus, in the model presented here, the reciprocity norm may indeed be a behavioral trait that a worker brings into the employment relationship from previous social experiences, but it may also be a trait that is learned in the workplace.

The model illustrates how an endogenous segmentation of firms with and without reciprocal employment relationships may develop based on heterogeneity of “returns to reciprocity.” Given that some firms have more to gain from engaging in a reciprocal relationship with their workers than others, they will also be more willing to incur the cost of developing such relationships. But since the level of reciprocity within the workforce is still determined in part by social forces, firms must also take the existing degree of this cultural norm into consideration in deciding whether or not to engage in a gift-exchange. This leads to a co-evolutionary dynamic between the economic institution of the firm and the cultural norm reciprocity. We derive the multiple equilibria produced by these dynamics and consider the welfare consequences of policies directed toward increasing reciprocity in the workplace.

In the next section of this paper, we present evidence of reciprocity in the employment relationship and attempt to draw some conclusions about when reciprocal relationships will emerge. Section II discusses the role of the firm in cultural evolution and presents the motivation for an evolutionary approach to reciprocity in the employment relationship. In Section III, we present the basic model of reciprocal employment relationships, and extend that model to allow reciprocating firms to screen imperfectly for reciprocating workers in Section IV. The welfare consequences of policies designed to increase the level of reciprocity are considered in Section V.

## **I. Evidence of Reciprocity**

### **Reciprocity in Economic Experiments**

The essential role that reciprocal gift-exchange has served in the social and economic structure of a variety of cultures has been well-documented by anthropologists (Mauss 1967, Schrift 1997, Komter 1996 and Baal 1975). The evident pervasiveness and functionality of the reciprocity norm—and its potential significance for a variety of economic interactions—has led to great interest among experimental economists in testing the nature of reciprocal behavior in the laboratory. Since the early work of Güth, et al. (1982), the “ultimatum game” has been run with a variety of experimental designs and many of these studies have found evidence of reciprocal behavior.<sup>4</sup> In the ultimatum game, subjects are put into pairs of “allocators” and “receivers.” The allocator is given a sum of money, and is instructed to split this money with the receivers. The allocator is free to determine the distribution of the money, subject only to the rule that he must give the receiver a non-zero portion of the pie. If the receiver accepts the offer, the money is distributed as the allocator proposed; if the receiver rejects the offer, both of the participants receive nothing.

In the great majority of ultimatum game experiments, the standard game theoretic prediction of purely strategically rational and selfish behavior (the allocator offering a

---

<sup>4</sup> See Fehr and Gächter (2000,1997), Guth (1995) and Roth and Kagel (1995) for a review of the evidence.

penny and the receiver accepting this distribution) can be rejected for a significant fraction of the participants. In fact, offers typically average around 20-40% of the pie. Not only do we frequently observe experimental subjects offering more than the minimum, but it is also often observed that if allocators do offer a very low amount, receivers will reject the distribution. In both cases (allocators offering more than a penny and receivers punishing very unequal distributions), subjects are exhibiting that they are willing to incur a cost to enforce a reciprocity norm. This result has been consistently generated in many versions of the ultimatum game.<sup>5</sup> Fehr and Gächter (2000,1997) survey ultimatum games in which the fraction of subjects who “exhibit reciprocal choices” is never below 40% and sometimes is above 60%. In many of these studies, a strong, statistically significant relationship is found between the level of the offer and the frequency of rejection. Note that these results have been observed even though partners are often anonymous and interactions are one-shot.

The ultimatum game has been run in many different forms, and the level of reciprocity is often quite sensitive to the experimental design. However, there are some general conclusions that can be drawn about reciprocal behavior from these experiments. First, some level of reciprocity almost always exists. In most of the ultimatum and dictator games that have been run, there is some combination of purely selfish actors (i.e. those who offer the minimum throughout the experiment) and some reciprocators. There are also typically some “strong” and some “weak” reciprocators—that is, some reciprocators who are willing to incur a cost to enforce the reciprocity norm even if they cannot benefit from this enforcement (e.g. if they punish on the last round of an experiment), and some reciprocators who will only act reciprocally if they can reasonably expect to gain from this behavior. Second, in most of these experiments the level of reciprocity increases with the frequency of interaction and the degree of “social history” shared by the subjects, decreases with the degree of anonymity and the use of market terminology, and is insensitive to the level of the pie.<sup>6</sup> Third, reciprocity is based on both the level and the perceived intentions of an offer.<sup>7</sup> The higher the offer in an ultimatum game, the higher the probability of acceptance, and a decrease in the perceived intentionality of the action—e.g. if the receiver does not know whether the allocator chose the distribution or if it was generated randomly—decreases the frequency of reciprocal responses.

### **Evidence of Reciprocity in the Employment Relationship**

A number of experiments have been designed to analyze the role of reciprocity in the employment relationship specifically. In a series of “double-auction” experiments, Fehr and Falk (1999), Falk and Gächter (1998) and Falk, et al. (1998) have firms and workers exchanging wages for effort when labor contracts are incomplete. Firms and workers are kept in separate rooms while they bargain over wages—firms can make wage offers that can be accepted or rejected by workers. Once one of the workers accepts a firm’s offer, the two are paired to “complete a contract.” If no workers accept the wage

---

<sup>5</sup> Dickinson (2000), Cameron (1999), Eckel and Grossman (1996), Forsythe, et al. (1994), Hoffman, et al. (1994).

<sup>6</sup> See Hoffman, et al. (1991, 1994), Berg, et al. (1995), Bolton and Zwick (1995), Cameron (1995), Abbink, et al. (1997) and Fehr and Gächter (2000).

<sup>7</sup> Falk and Fischbacher (1998), Bount (1995), and Kahneman, et al. (1986).

offers made by the firms, they can make counter-offers, which, if accepted by the firms, will also initiate a pairing to complete a contract. Once workers and firms are paired, workers must choose the level of effort they will give the firm. The firms' payoffs are increasing in the level of effort and decreasing in the wage, while the opposite holds for the worker. All of these interactions are anonymous and one-shot.

The null hypothesis that is tested in each of these experiments is the game theoretic prediction that workers have no incentive to give above-minimum effort and, thus, firms, since they expect workers to give minimum effort, will offer the wage associated with that minimum effort level. Furthermore, it will always be in the worker's best interest to offer minimum effort so, it is predicted, higher effort will not be elicited by higher wages or vice versa. What has been repeatedly observed, however, is that 1) average wage offers exceed the minimum level (i.e. the level associated with minimum effort), 2) average effort levels exceed the minimum level and 3) there is a positive correlation between wages and effort. As with most of the ultimatum game experiments, the authors consistently observe some fraction of firms and workers that do not reciprocate, offering minimum wage and effort levels throughout the experiment.

Falk and Gächter (1998) design a version of this double auction experiment in which they test to see whether wages will fall to the minimum level if there is a surplus of workers over firms (even though workers still have the ability to choose effort levels). The workers who do not complete a contract will not receive any money in that round and thus it is in each worker's interest to be paired with a firm, even if they are receiving the minimum wage. The authors find that wages and effort levels are very close to the levels observed in the experiments with an equal number of workers and firms, even though workers try to underbid one another (wages converge to about 40-60% of the highest level, and effort levels converge to about 20-40% of the highest level). The authors argue that firms continue to offer above-minimum wages, even though they know that they could hire the under-bidders, because they (correctly) predict that this will induce the workers with which they are paired to give high effort.

In his extensive survey of Northeastern U.S. firms during the early 1990s recession, Bewley (2000) finds evidence that firms do indeed behave as the "firms" in the Falk and Gächter experiment. Bewley talks with a number of firms who tell him that the reason they are reluctant to cut wages during a recession is because it will damage work morale and allegiance to the organization. Also supporting the conclusions of Falk and Gächter, Bewley finds that firms do not reduce wages, even when there is a surplus of workers who are willing to work at lower wages than the current employees. Many of the firms argue that cutting wages will lead to reduced productivity and higher turnover, and hiring under-bidders will hurt worker cohesion. Further, the firms that Bewley surveys who did in fact cut wages during the recession, reported that they experienced productivity losses. This suggests that a number of firms not only *expect* to have reciprocal relationships with their employees, but believe that they can influence whether and how these relationships will develop.

Kahneman, et al. (1986) also find that firms are concerned about fairness in their wage-setting behavior. They find that workers are more willing to accept wage cuts when a firm is experiencing a decline in profits than when it is responding to changes in labor supply. Not only is this evidence of reciprocal employment relationships in some firms, but it is also evidence that the level of worker reciprocity is sensitive to the perceived

intentionality of the firm's behavior (Falk and Fischbacher 1998). That is, worker reciprocity is not only responsive to changes in wages, but to whether or not the firm's behavior is perceived as chosen (responding to labor supply) or forced (experiencing a decline in profits).

### **Workplace Characteristics Influencing Reciprocity**

From the evidence of reciprocity discussed above, there are certain hypotheses we can make about the characteristics of a work environment that might lead to reciprocal employment relationships. First, of course, is the compensation mechanism.<sup>8</sup> It is clear that, in general, higher levels of compensation improve reciprocity. However, there are a number of factors influencing whether workers perceive a wage as a "gift." Information that the worker uses in perceiving the intentionality of the gift may not only be taken from the *cause* of the change in wages (as in Kahneman, et al. 1986), but from the degree to which the gift is contingent on performance. Corresponding to the experimental evidence that the level of reciprocity will increase when individuals are able to directly reward those who treat them kindly and directly punish those who did not, we would expect "discretionary effort" (the worker's gift) to be more responsive to, for instance, an individual bonus, than to an increase in the average level of pay of all employees. Workers may also judge the fairness of a given level of compensation by comparing it to other workers in the firm, in their occupation, etc.<sup>9</sup>

Compensation mechanisms are not the only characteristics of the workplace that may influence reciprocity. The method and degree of supervision in a workplace may have also have implications for the development of the reciprocal employment relationship.<sup>10</sup> This is because supervision, by providing a direct extrinsic incentive to increase effort, may "crowd out" the worker's intrinsic motivation to perform the task (or offer a gift).<sup>11</sup> Another interpretation of the disincentive effects of supervision is that it provides a signal to the workers that they are not trusted and thus that the employer is not interested in initiating a gift exchange. Reciprocal employment relationships may thus be more likely to arise in firms that use "carrots" to motivate workers, rather than "sticks."

One occupational characteristic cited by organizational psychologists as contributing to the level of reciprocity in the firm is the scope for employee participation in job design.<sup>12</sup> Although there have been a number of arguments put forward for why this might be the case, most of them are centered on the idea that, beyond some minimum standard set by the firm, effort can vary dramatically with the degree of employee motivation. Motivation, it is argued, is largely the result of employee satisfaction, which is, in turn, positively influenced by the scope for employee participation.

Although the evidence of the impact of employee participation on productivity has been mixed in the past,<sup>13</sup> some recent studies have found strong productivity

---

<sup>8</sup> See Guzzo and Katzell (1987) for a review of the psychological evidence of the effects of economic incentives on productivity.

<sup>9</sup> See, e.g. Akerlof and Yellen (1990).

<sup>10</sup> Frey (1993, 1994, 1997), Chang and Lai (1999), Drago and Perlman (1989) and Guzzo and Katzell (1987).

<sup>11</sup> Frey (2001, 1997), Kreps (1997), Deci (1971) and Lepper, et al. (1973).

<sup>12</sup> Guzzo and Katzell (1987), Miller and Monge (1986), Locke and Schweiger (1979) and Herzberg (1968).

<sup>13</sup> See Miller and Monge (1986), Locke and Schweiger (1979) and Herzberg (1968) for reviews of the early evidence.

increases from the introduction of employee participation. For example, a number of recent studies of high performance work organizations (HPWO) have shown that methods of work organization that involve workers in product design, engineering and quality control have had a significant positive impact on productivity.<sup>14</sup> A growing number of firms have adopted methods of HPWO (e.g. team production, job rotation, quality circles, etc.) in the past decade,<sup>15</sup> and some general characteristics of HPWO can be observed across firms which suggest that these firms are capturing the gains to reciprocal employment relationships. In addition to providing training and scope for employee participation, most HPWOs include compensation mechanisms which tie pay directly to effort (such as profit sharing and “pay for knowledge”).<sup>16</sup> Furthermore, firms that adopt HPWO will also frequently screen intensively for workers with “personality traits needed for cooperative team environments” (Ichniowski, et al. 1997). Firms that have adopted HPWO seem to be depending, to some extent, on the level of reciprocity of their workforce, since the profitability of methods of work organization such as lean production, just-in-time manufacturing and quality circles depend on the “discretionary effort” (MacDuffie 1995) that employees put into problem-solving.

High performance work organizations are not the only types of firms which appear to be engaging in reciprocal employment relationships with their employees. As the evidence discussed above suggests, the gift-exchange that Akerlof (1982) highlighted among the cash posters seems to appear across many different types of occupations, industries and work environments. What seems to be workers offering “discretionary effort,” if not ubiquitous, is certainly quite common. Although all of the hypothesized influences on reciprocal employment relationships discussed in this section will not be explored in the model presented here, the model is intended to provide a broad framework for analyzing the role of the firm in the adoption of this social norm. In presenting an analysis of how economic and cultural institutions may co-evolve, we also offer an explanation for why reciprocal employment relationships may develop in certain types of firms and not others. In the next section, we provide some background for the model with a discussion of cultural evolution and the potential role of the firm in the cultural evolutionary process. We also provide motivation for an evolutionary game theoretic approach to reciprocal employment relationships.

## **II. Cultural Evolution and the Reciprocal Employment Relationship**

### **The Role of the Firm in Cultural Inheritance**

There are two fundamental dimensions to the process of cultural inheritance—that is, how a cultural trait such as reciprocity is adopted and passed on within society (Boyd and Richerson 1985, Cavalli-Sforza and Feldman 1982). The first, *cultural transmission*, is the channel through which an individual or a group actually adopts the cultural trait. For example, one might begin to behave according to a particular cultural norm through learning, imitation or imprinting (unconscious adoption of a cultural trait). The second

---

<sup>14</sup> Ichniowski, et al. (1997), MacDuffie (1995), Batt and Appelbaum (1994).

<sup>15</sup> Osterman (2000, 1999) and Batt and Appelbaum (1994).

<sup>16</sup> In fact, it has been argued that a firm will only realize the productivity gains to switching to HPWO if these methods of work organization and compensation are adopted together. See, e.g. Ichniowski, et al. (1997), MacDuffie (1995), Baker, et al. (1994).

main element of the cultural inheritance process is *cultural selection*—that is, the mechanisms influencing which traits will be adopted, how those traits will be diffused throughout society, and whether such traits will be sustained.

In the model of workplace reciprocity presented here, we consider the potential role of the firm in the cultural selection process. That is, we assume that there are various methods by which the reciprocity trait may be passed on (transmitted) from one individual to another, and consider how the firm might influence who adopts this cultural trait within society. The firm plays a role in cultural inheritance on our model by encouraging or discouraging the adoption of the reciprocity norm by assigning costs or benefits to reciprocal behavior. The structure of social interactions will affect how likely it will be that individuals will learn about the reciprocity norm and will influence the diffusion of the norm throughout society. We use an evolutionary game theory model to analyze the role of the firm in shaping the evolution of the reciprocity norm.

In the model presented below, the motivation for behavior is situated in a social context. This seems to be a more appropriate context for motivating norm-guided behavior than strategic decision-making. There is good evidence that people do not act reciprocally for purely strategic reasons—that is, that reciprocal behavior is often exhibited in circumstances in which the person does not gain materially from the behavior.<sup>17</sup> While one approach to analyzing such behavior may be to simply assume that some people derive utility from reciprocity and choose the optimal behavior on the basis of this, this approach is problematic for two reasons. First, it does not allow us to say anything meaningful about who will exhibit this behavior and when, and thus theoretical frameworks built on this assumption will not yield predictions about how reciprocity could be stimulated with policy. Of course, it is irrelevant that a strategic analysis of reciprocity will yield few testable hypotheses if this approach is indeed correct. But the experimental evidence discussed in the previous section suggests that it is not. This evidence much more strongly supports the notion that reciprocal behavior, rather than being the outcome of a strategic decision, is learned through a series of social and economic interactions. Far from implying that such behavior is hardwired and thus applied sub-optimally in many contexts, this latter view of reciprocal behavior implies that people learn when reciprocity is appropriate by experiencing the rewards and punishment to such behavior in a variety of circumstances. This explanation for reciprocal behavior (the cultural evolutionary perspective) implies that, if we want to understand when and why people behave reciprocally, we must understand how the behavior is learned and adopted in various contexts.

Norm-guided is often taken to mean hardwired—that is, that people who “possess” this norm will be motivated by it in every context. But theories of cultural evolution provide a framework for understanding how norm-guided behavior can be flexible, because such behavior can be learned through social and economic experiences. The possibility that optimal behavior requires learning through social interaction is the second reason why the strategic approach to reciprocity may be an inappropriate framework for analysis. The standard assumption that people can immediately learn how to behave optimally may actually preclude the generation of realistic analytical results. In the model presented here, the possibility of a stable internal equilibrium in which some

---

<sup>17</sup> For example, leaving a tip in exchange for good service even when one does not intend to return to the restaurant, or punishing in the last round of the ultimatum game.

workers and firms behave reciprocally and some do not depends on people learning slowly, through social interactions, what the correct behavior is for a given change in the environment. If people knew exactly when and how to switch behavior, as they do in the standard game theoretic approach, the population would simply jump back and forth between behavioral regimes. The population would either never converge to equilibrium or would unravel into equilibria in which everyone or no one reciprocates. It is often the case that an evolutionary analysis of a game yields the same pure-strategy equilibria as the standard game, and also yields an internal equilibrium corresponding to the mixed strategy. If we believe that we do indeed observe some firms and workers who engage in reciprocal relationships and some firms and workers who do not—and if this is not the result of individuals playing mixed strategies—then our model predicts that this is because people do not learn immediately what behavior is best for them all of the time.

While models of rational choice can often yield similar *outcomes* to a cultural evolutionary model, the representation of how the behavior is *motivated* in these two approaches can have very different policy implications. In order to know how we might encourage reciprocal behavior—or whether it might be influenced by the existing arrangement of economic institutions—we must first know whether the assumptions we are making about the reasons for reciprocal behavior will yield accurate predictions. Similarly, in order to understand the potential welfare effects of changes in policy, we may need to understand the motivations for reciprocal behavior. In the next section, a theoretical approach to reciprocity in the employment relationship is developed. Section IV presents an extension to the basic model, in which firms that want to reciprocate are able to screen imperfectly for reciprocators. In the last section, the model is used to make predictions about the implications of changes in economic policy on reciprocal behavior and welfare. We hope to test these predictions in future work.

### **III. A Basic Model of Reciprocity in the Employment Relationship**

#### **Introduction**

We begin with a simple model of the employment relationship, conceived of as a coordination game between workers and firms. It is assumed that workers are heterogeneous in their behavior in this game, with some fraction of them behaving according to a reciprocity norm and the rest behaving non-reciprocally.<sup>18</sup> Firms can also act reciprocally or not, but are assumed to be able to choose the optimal behavior, whereas workers—whose behavior is norm-guided—always behave according to the norm they possess. This does not mean that worker behavior is unchanging, but rather that a worker's optimal behavior must be *learned*, both through the worker's interaction with different types of firms and through her social interactions with other workers. The reciprocity norm thus evolves in the workforce through the differential replication process of cultural selection and transmission discussed above.

In the model presented below, the cultural selection pressures acting on the reciprocity norm are shaped in part by the structure of the firms. That is, it is the costs or benefits to worker reciprocity within a particular type of firm that affects whether the worker will continue to behave in this way. Intuitively, if reciprocal behavior is rewarded

---

<sup>18</sup> The experimental tests of reciprocity discussed in the previous section provide strong and consistent evidence that such heterogeneity in the propensity to reciprocate exists in society.

in the workplace—if gains to reciprocal behavior are perceived—the behavior will be encouraged. However, a worker’s behavior is not influenced exclusively by whether such behavior has been rewarded or punished in previous work environments, but also by social interactions outside of the workplace, where he can learn trends in firm behavior and the experiences of his peers who have behaved in different ways. This latter influence on behavior, which is intended to represent the cultural transmission process, is not modeled precisely—that is, a *particular* method of cultural transmission is not assumed—but is rather represented generally as the result of structured social interactions. This is the method employed by Bowles and Gintis (1998) in their model of cooperation within small communities.

One can imagine the reciprocity and non-reciprocity traits possessed by the workers as determining the level of “discretionary effort” that they give to the firm. For instance, the worker could produce a higher quantity or be more careful in determining quality than is minimally enforceable, as with the cash posting women who motivate the Akerlof gift-exchange model (Akerlof 1982). Employees who are working in a high performance work organization also have (within reasonable bounds) discretion over how much effort they will put toward identifying the source of a product defect, suggesting feasible quality improvements, etc. Firms that have organized their workplace to take advantage of, for instance, team production, “just in time” manufacturing or lean production, rely on the motivation of their employees to analyze production problems, to offer suggestions for improvement and to work cooperatively with other employees. It is in this sense that “discretionary effort” can be fundamental to profitability.

In deciding whether to behave reciprocally or not with its worker, the firm may be deciding whether to switch production over to HPWO. The firm may not be deciding to transform its *entire* workplace, but may choose to adopt only some elements of the HPWO (e.g. it may only introduce quality circles without introducing job rotation) (Appelbaum and Batt 1994). It is typically the case that firms that switch (at least partially) to HPWO, also switch the form of compensation to a method that has been shown to be effective in this context, such as profit-sharing, “pay-for-knowledge” incentive schemes, etc. (Ichniowski et. al. 1997, MacDuffie 1995, and Kandel and Lazear 1992).

In the model presented below, it is assumed that firms will vary in their “returns to reciprocity,” and thus the value of a reciprocating worker will differ across firms. A particular firm’s returns to reciprocity could be determined by their production technology and product market. How many firms value and reward reciprocal behavior by their workers will influence how many workers adopt reciprocal behavior which will, in turn, influence how profitable it is for a firm to behave reciprocally itself. The basic model presented below is an analysis of these co-evolutionary dynamics. The model is then extended to allow firms to screen for reciprocal characteristics in their workers. As is common in models with strategic complementarities, the model below will yield multiple equilibria which can be Pareto ranked by analyzing the welfare consequences of the various external effects (Cooper and John 1988).

### **Setup**

Assume that there are two large populations, one of workers ( $w$ ) and one of firms ( $f$ ), each with unit measure. Workers can be characterized as possessing one of two *traits*,

$t = r, n$ , which determine their behavior in a coordination game with firms. When paired with a firm to engage in an employment relationship, workers who possess the  $r$  trait will reciprocate and those which possess the  $n$  trait will not reciprocate. Firms will choose the strategy (reciprocate or don't reciprocate) which yields the highest expected payoffs, as described below.

Call  $p_i^t$  the fraction of reciprocators in sub-population  $i$  ( $i = w, f$ ) at time  $t$ . Workers and firms are randomly selected from the population to engage in a one-shot employment relationship, where they exchange effort for wages.<sup>19</sup> This game is repeated continually with new worker-firm pairings in each round.

Call  $p_i^{tn}$  the payoff to an  $i$ -type with trait  $t$  from engaging in an employment relationship with a not- $i$  person of trait  $u$  ( $u = r, n$ ). The firm's payoff to mutual reciprocation is further subscripted  $p_{ff}^{rr}$  to denote differences in this payoff across firms. All other payoffs are constant. The normal form of the one-shot game, with the payoff to workers shown first is:

		<b>Firms</b>	
		Reciprocate	Don't Reciprocate
<b>Workers</b>	Reciprocate	$p_w^{rr}, p_{ff}^{rr}$	$p_w^{rn}, p_f^{nr}$
	Don't Reciprocate	$p_w^{nr}, p_f^{rn}$	$p_w^{nn}, p_f^{nn}$

It is assumed that the following relationship holds across payoffs:

$$(1) \quad p_i^{rr} > p_i^{nr} > p_i^{nn} > p_i^{rn}$$

This ordering of payoffs implies that the workers and firms are engaging in a "coordination game" with two pure strategy Nash equilibria (mutual reciprocation and mutual non-reciprocation), with mutual reciprocation Pareto dominating mutual non-reciprocation.

---

<sup>19</sup> Even though the employment relationship is modeled as one-shot, there is still scope for reciprocal interaction. For example, since the length of a period has not been specified, the payoffs from a particular game could be assumed to be a summation of a number of interactions between the firm and worker.

We assume that firms can be ordered along a continuum of “returns to reciprocity”.<sup>20</sup> The density of firms along the continuum of returns to reciprocity is denoted  $f(p_{ff}^{rr})$  with support  $(p_f^{nr}, \infty)$ .

Expected payoff functions for workers and firms of each trait can also be defined. Note that, since players are assumed to be drawn randomly from a large population, the probability of interacting with another player of a given trait is equal to the fraction of the sub-population that possesses that trait. When screening is introduced in the expanded model below, it will no longer be the case that the probability of interaction is given by the frequency of traits in the population. Call  $Ep_i^t(t)$  the expected payoff to an  $i$ -type player with trait  $t$  to interacting with a  $-i$  player with either trait  $\tau$  or trait  $-\tau$  at time  $t$ .

$$(2) \quad Ep_i^t(t) = p_{-i}^t p_{tr}^i + (1 - p_{-i}^t) p_{tn}^i$$

### Replication

The differential replication process within the worker sub-population takes place as follows:<sup>21</sup> Assume that in every period, each worker with probability  $d_1 > 0$  decides to evaluate the benefit to changing traits. If that worker encounters a worker of the other trait, he will compare expected payoffs, and if the expected payoff to being the other trait is higher, he will switch with a probability proportional to  $(Ep_i^t(r) - Ep_i^t(n))$ , with proportionality factor  $d_2 > 0$ .<sup>22</sup> Given these assumptions about social interaction and rules for the adoption of traits, the evolution of the reciprocity trait within the worker population can be represented with the following replicator function:<sup>23</sup>

$$(3) \quad R_r^t = d_1 d_2 p_w^t (1 - p_w^t) [Ep_w^t(r) - Ep_w^t(n)]$$

The replicator function represents the change in the level of reciprocity within the worker population between period  $t$  and period  $(t + 1)$ . It is positive—that is, the  $r$  trait is

<sup>20</sup> Since the replicator function described below depends on workers’ expected payoffs, allowing for heterogeneity of payoffs across workers would not change the results. Homogeneity of worker payoffs is thus assumed for simplicity.

<sup>21</sup> The social mechanism underlying the differential replication process in this model is based on Bowles and Gintis (1998).

<sup>22</sup> There are a number of ways in which this process of cultural learning among workers might be interpreted. For example, workers within a particular firm, industry, neighborhood, etc. may discuss their experiences offering various levels of effort. Of course, in each of these scenarios, the worker would face a different probability of encountering a worker with the other trait (e.g. it may be much higher within one’s firm than within one’s industry). This is the basis of differential replication models of structured populations (Boyd and Richerson 1985). The differential replication model presented here (in which the probability of interaction is simply determined by the frequency of traits within the population) is the most general version of a structured population model, and introducing group-specific probabilities of interaction would not alter the general results of the model for the society, but may affect the likelihood that a given group will move toward a particular equilibrium. I also leave open the method of cultural transmission (i.e. how a worker actually acquires the reciprocity or non-reciprocity trait).

<sup>23</sup> See Bowles and Gintis (1998) and Hirshleifer and Coll (1988) for applications of similar replicator functions.

increasing in the population—if the expected payoff to being an  $r$  type is higher than the expected payoff to being a  $y$  type. This yields the following expression for  $p_w^t$  :

$$(4) \quad p_w^t = p_w^{t-1} + R_r^{t-1}$$

It is assumed that firms choose to reciprocate or not reciprocate by choosing the trait which yields the maximum expected payoff. The number of firms who are reciprocating in a given period will depend both on the number of reciprocating workers and on the distribution of firms along the continuum of returns to reciprocity. The expression for  $p_f^t$  can be derived simply by finding the minimum level of  $p_{ff}^{rr}$  a firm must possess in order for  $Ep_{ff}^t(r) > Ep_{ff}^t(n)$  for any given  $p_w^t$ . That is, those firms will be reciprocators for which  $p_{ff}^{rr} \geq p_f^{nr} + \frac{(1-p_w^t)}{p_w^t} [p_f^{nn} - p_f^{rn}]$ .

$$(5) \quad p_f^t = \int_{p_f^{nr} + \frac{(1-p_w^t)}{p_w^t} [p_f^{nn} - p_f^{rn}]}^{\infty} f(p_f^{rr}) dp_f^{rr}$$

### Equilibrium

Equilibrium  $(p_w^*, p_f^*)$  is defined as stationarity of traits, occurring when the replicator function is equal to zero. It is easy to see that equilibrium occurs at (0,0) and (1,1) and that the internal equilibrium  $(p_w^{\text{int}}, p_f^{\text{int}})$  occurs when the expected payoffs across traits in the worker population are equal:

$$(6) \quad Ep_w(r) = Ep_w(n)$$

Recall that the workers' expected payoffs are functions of level of reciprocity in the *firm* population, and thus a particular value  $p_f$  will solve Equation (6). Substituting Equation (2) for both sides of Equation (6) and solving for  $p_f^{\text{int}}$  yields:

$$(7) \quad p_f^{\text{int}} = \frac{p_w^{nn} - p_w^{rn}}{p_w^{nn} - p_w^{rn} + p_w^{rr} - p_w^{nr}}$$

Notice that  $0 < p_f^{\text{int}} < 1$  by the conditions specified for the payoffs in Equation (1).

Assuming continuity of  $f(p_{ff}^{rr})$  an equilibrium level of reciprocity in the worker population  $(p_w^{\text{int}})$  must solve:

$$(8) \quad p_f^{\text{int}} = \int_{p_f^{nr} + \frac{(1-p_w^{\text{int}})}{p_w^{\text{int}}} [p_f^{nn} - p_f^{rn}]}^{\infty} f(p_f^{rr}) dp_f^{rr} = \frac{p_w^{nn} - p_w^{rn}}{p_w^{nn} - p_w^{rn} + p_w^{rr} - p_w^{nr}}$$

Since the value of  $p_f^t$  given by Equation (5) goes one to zero as  $p_w^t$  goes from one to zero, as long as  $f$  has positive density in the neighborhood of the equilibrium, there will exist a unique internal equilibrium level of reciprocity in the worker population  $p_w^{\text{int}}$ , corresponding to the equilibrium level of reciprocity in the firm population  $p_f^{\text{int}}$ .

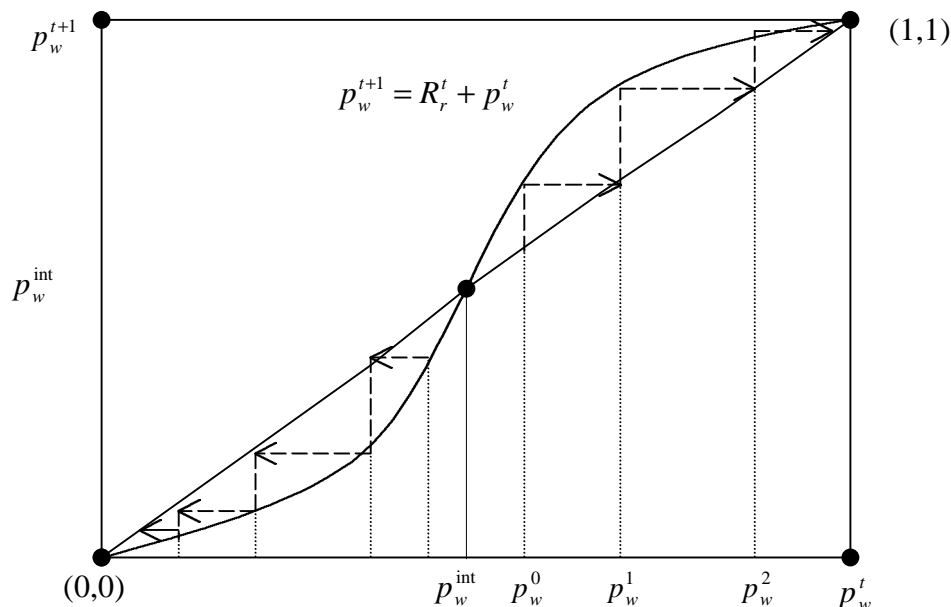
### **Stability**

One can see a few things about the dynamics of this system simply by inspection of (3), (4) and (5). Notice that  $p_f$  is an increasing function of  $p_w$ , and that  $p_w$  is increasing (i.e.  $R_r > 0$ ) when  $p_f > p_f^{\text{int}}$  and decreasing when  $p_f < p_f^{\text{int}}$ .

The dynamics of this system can be seen clearly with a diagram of Equation 4:

**Figure 1**

**Dynamics in the Worker Population (Without Screening)**



In Figure 1, the  $p_w$  function crosses the 45 degree line when the replicator function is equal to zero, is above this line for  $R_r > 0$ , and below it for  $R_r < 0$ . Because the  $p_w$  function crosses the 45 degree line from below at the internal equilibrium, this equilibrium is unstable. Any small perturbations from  $(p_w^{int}, p_f^{int})$  will cause the population to unravel to one of the corner equilibria. It is intuitive why the system exhibits these dynamics: First, since firms are behaving rationally, the level of reciprocity in the firm population moves in the same direction as the level of reciprocity in the worker population, because, for example, an increase in  $p_w$  increases the expected payoffs to being a reciprocating firm, and thus decreases the minimum level of  $p_{ff}^{rr}$  required for a firm's expected payoffs to reciprocating to be higher than its expected payoffs to not reciprocating. That is, more firms along the continuum of returns to reciprocity will find it in their interest to be a reciprocator for an increase in  $p_w$ , and more firms will switch to non-reciprocation for a decrease in  $p_w$ . The workers' behavior, on the other hand, depends entirely on which side of the internal equilibrium the population is, and  $p_w$  will be declining if  $p_w < p_w^{int}$  and increasing if  $p_w > p_w^{int}$ .

The prediction that this model yields about the impossibility of maintaining a stable internal level of firms which choose to reciprocate with their workers seems unrealistic. We do indeed observe some fraction of firms, in the U.S. for instance, which engage in reciprocal interactions with their workforces, and organize production and

compensation to take advantage of this cooperative relationship.<sup>24</sup> Similarly, we observe workers who are willing to engage in these relationships as well.

So must it really be the case, as the model predicts, that if there are relatively few workers willing to engage in reciprocal employment relationships (low  $p_w$ ), that it will be in the interest of very few firms to have these relationships as well? It is clear that, even in this simple model, there are firms who want to be reciprocators even when the majority of the workers in the labor force are not. This is because of the particularly high returns to reciprocity within these firms. Think, however, of the change in the level of  $p_f$  when  $p_w$  declines—firms who are on the margin between the reciprocation and non-reciprocation regimes will switch to being non-reciprocators. That is, the firms are primarily constrained in choosing optimal behavior by the number of reciprocating workers that apply for its jobs. In reality, however, most firms' production decisions do not seem to be so sensitive to the number of cooperative workers they can find. Rather, if the returns to a firm's cooperative behavior are decreasing because most of the workers it finds are not willing to participate, it is more likely that the firm will choose to screen more intensively for these types of workers than switch to a non-reciprocal method of production. In fact, many U.S. firms that have switched to HPWO extensively screen their workers (often for up to a year) for what they describe as "cooperative" personality traits (Ichniowski, et al. 1997).

In the elaborated model presented below, I allow for the possibility of screening for reciprocating workers. This will become an optimal strategy for some firms, depending on their particular returns to reciprocity (i.e. how costly it is for them to be employing a non-reciprocator) and the level of reciprocity in the worker population. It will be shown that there are also firms with sufficiently low returns to reciprocity, that it will never be in their interest to pay the screening cost to find a reciprocating worker. Below it is illustrated that, once the possibility of using an imperfect screening mechanism is introduced, there is the possibility of the existence of at least one locally stable internal equilibrium.

#### IV. The Reciprocal Employment Relationship with Screening

##### Setup

Now assume that firms can choose one of three strategies for their behavior in the employment relationship—they can reciprocate ( $r$ ) or not reciprocate ( $n$ ) as above, or they can screen ( $s$ ). It is assumed that the screening technology improves the probability that a firm will hire a reciprocating worker (i.e. the screening firm will face a probability of finding a reciprocating worker that is higher than the frequency of reciprocating workers in the population). For simplicity, it is assumed that all firms face an identical screening cost ( $c$ ).

Let  $S(c, p_w^t)$  be a continuously differentiable function which yields the fraction of firms that choose to screen in period  $t$ , with  $S_1 < 0$  and  $S_2 > 0$ . The fraction of firms that will want to screen is a function of the cost of screening and the fraction of

---

<sup>24</sup> See Freeman, et al. (2000) and Osterman (2000) for estimates of the fraction of firms adopting HPWO.

reciprocating workers in the population. Call  $C$  the fraction of firms which are reciprocators, but do not screen and thus:  $p_f^t = S^t + C^t$ .

Call  $q_{f\tau}^t$  ( $\tau = s, r, n$ ) the probability that a firm that chooses strategy  $\tau$  will find a reciprocating worker in period  $t$ .  $q_{fs}^t(p_w^t, S^t)$  will be increasing function of  $p_w^t$  and a decreasing function of the number of firms which are screening ( $S^t$ ). Since screening increases the probability of finding a reciprocating worker,  $q_{fs}^t > q_{f,-s}^t$ . Note also that:

$$(9) \quad q_{fr}^t = q_{fn}^t = \frac{p_w^t - S^t q_{fs}^t}{1 - S^t}$$

Equation (9) simply says that the probability of either a reciprocating or a non-reciprocating firm that does not screen being paired with a reciprocating worker is equal to the ratio of reciprocating workers which are not hired by screeners to the overall number of workers which have not been hired by screeners. Expected payoff functions for the non-screening firms are similar to those specified in the basic model (Equation (2)), except that the probability of encountering a worker of a given trait is now defined by Equation (9). The expected payoff to a screening firm is given by:

$$(10) \quad Ep^t(s) = q_{fs}^t p_{ff}^{rr} + (1 - q_{fs}^t) p_f^m - c$$

As before, firms choose their strategy to maximize expected payoffs.

Now consider the effect that screening has on the worker population. The probability that a reciprocating worker will engage in an employment relationship with a reciprocating firm in time  $t$ , denoted  $q_{wr}^t$  is now:

$$(11) \quad q_{wr}^t = \frac{S^t q_{fs}^t + (p_f^t - S^t) \frac{(p_w^t - S^t q_{fs}^t)}{(1 - S^t)}}{p_w^t} = \frac{S^t q_{fs}^t (1 - p_f^t) + p_w^t (p_f^t - S^t)}{p_w^t (1 - S^t)}$$

In Equation (11) the probability that a reciprocating worker will be matched with a reciprocating firm is equal to the ratio of reciprocating workers who are matched with either screening firms ( $S^t q_{fs}^t$ ) or non-screening reciprocating firms

$((p_f^t - S^t) \frac{(p_w^t - S^t q_{fs}^t)}{(1 - S^t)})$  to the total number of reciprocating workers in the population

$(p_w^t)$ . The probability that a non-reciprocating worker will engage in an employment relationship with a reciprocating firm will be:

$$(12) \quad q_{wn}^t = \frac{S^t(1 - q_{fs}^t) + (p_f^t - S^t) \left[ 1 - \frac{(p_w^t - S^t q_{fs}^t)}{(1 - S^t)} \right]}{(1 - p_w^t)}$$

$$= \frac{S^t(1 - q_{fs}^t)(1 - S^t) - (p_f^t - S^t) [S^t(1 - q_{fs}^t) - (1 - p_w^t)]}{(1 - S^t)(1 - p_w^t)}$$

Note that if there are no screening firms ( $S = 0$ ), then both  $q_{wr}$  and  $q_{wn}$  are equal to the fraction of reciprocators in the firm population ( $q_{wr} = p_f$ ). If all of the firms are reciprocating screeners ( $S = p_f = 1$ ), then both  $q_{wr}$  and  $q_{wn}$  are equal to 1. We can also write:

$$(13) \quad q_{wn}^t = \frac{p_f^t - p_w^t q_{wr}^t}{1 - p_w^t}$$

Expected payoffs for workers have the same form as in the basic model (Equation (2)), but the probabilities of interaction will be given by (11) and (12).

The replicator function will also have the same form as above (Equation (3)), except that workers will now be comparing expected payoffs with the new probabilities of interaction.

### Equilibria

Once again, there are equilibria at (0,0) and (1,1), and internal equilibria which are defined by the level of reciprocity in the *firm* population that equates workers' expected payoffs across traits. Using (11), (12) and (13), we can find the expression for  $p_f^{\text{int}}$  that solves  $Ep_w(r) = Ep_w(n)$ :

$$(15) \quad p_f^{\text{int}} = \frac{S \left\{ (p_w - q_{fs}) [p_w (p_w^{nr} - p_w^{nm}) + (1 - p_w) (p_w^{rr} - p_w^{rn})] \right\} + (1 - S) p_w (1 - p_w) [p_w^{nm} - p_w^{rn}]}{(p_w - S q_{fs}) \left\{ p_w [p_w^{nr} - p_w^{nm}] + (1 - p_w) [p_w^{rr} - p_w^{rn}] \right\} - p_w (1 - S) [p_w^{nr} - p_w^{nm}]}$$

The level of reciprocity in the worker population at an internal equilibrium ( $p_w^{\text{int}}$ ) is defined implicitly by (15) and need not be unique. In fact, in the case that is analyzed below—the case in which there is the possibility of an internal locally stable equilibrium— $p_w^{\text{int}}$  need not be unique. Note that at  $S = 0$ ,  $p_f^{\text{int}}$  is equal to its value in the basic model (Equation (7)).

To understand the role that screening plays in the dynamics of the model and in creating the possibility of a locally stable internal equilibrium, consider the following simple example.<sup>25</sup> Suppose that, instead of a continuous positive density of firms, there are only two types of firms in this economy, one type of firm with returns to reciprocity  $p_1^{rr}$  and another with  $p_2^{rr}$ . It is assumed that  $p_1^{rr}$  is sufficiently low (see Appendix) that there are no values of  $p_w$  such the firm 1 types will want to screen. Firm 1 types will

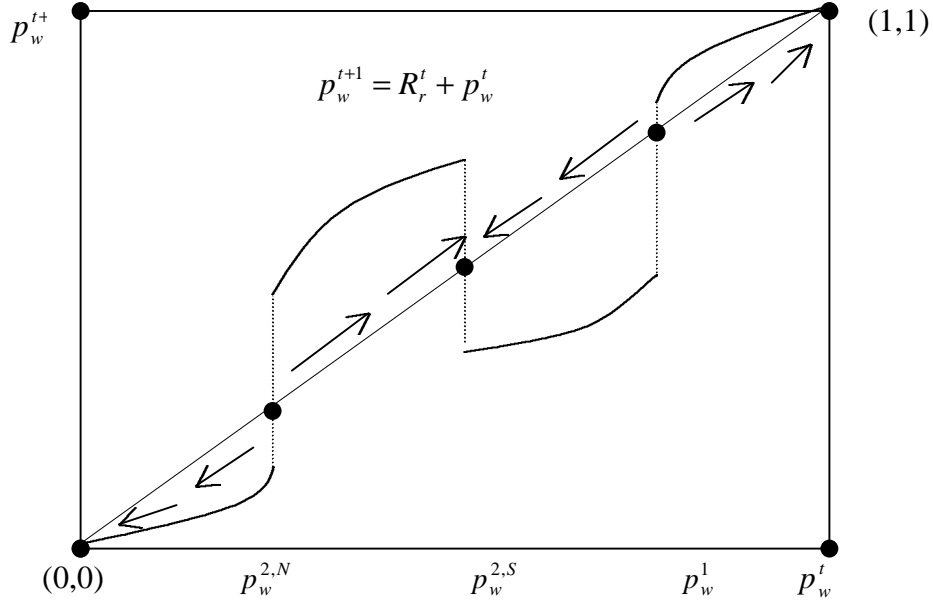
<sup>25</sup> A more detailed exposition and a formal proof of local stability is provided in the Appendix.

switch from reciprocity to non-reciprocity at some critical level of  $p_w$ , denoted  $p_w^1$ . For levels of  $p_w > p_w^1$ , firm 1 types will be reciprocators, and for  $p_w < p_w^1$  they will be non-reciprocators. Suppose that  $p_2^{rr}$  is sufficiently high so that it is possible for firm 2 types to be reciprocators, non-reciprocators or screeners, for different values of  $p_w$ . Call  $p_w^{2,S}$  the level of reciprocity in the firm population, above which the firm 2 types will be reciprocators and below which they will be screeners. Call  $p_w^{2,N}$  the critical value of  $p_w$ , above which the firm 2 types will be screeners and below which they will be non-reciprocators. Thus, the firm 2 types will reciprocate for  $1 \leq p_w < p_w^{2,S}$ , will screen for  $p_w^{2,S} \leq p_w < p_w^{2,N}$  and will not reciprocate for  $0 \leq p_w \leq p_w^{2,N}$ .

With firms of only these two types, the following dynamics will be observed in the worker population (derived from Equation (4)):

**Figure 2**

**Dynamics in the Worker Population with Screening and Only Two Types of Firms**



Type 1:	-----NR----- -----R-----
Type 2:	---NR--- ---S--- ---R---

Recall that the population is in equilibrium when the  $p_w^{t+1}$  function crosses the 45 degree line (i.e. the replicator function ( $R_r^t$ ) is equal to zero). Note also that the function is above the 45 degree line when the replicator function is positive—that is, when the expected payoffs to being a reciprocating worker are higher than to being a non-reciprocating worker—and below the 45 degree line when the replicator function is negative.

In the region  $1 \geq p_w^t > p_w^1$  both types of firms will want to reciprocate. At that level of worker reciprocity, the expected payoffs to the reciprocation strategy are higher than the expected payoffs to non-reciprocation for Type 1 firms and higher than both the expected payoffs to non-reciprocation and screening for Type 2 firms. When all firms are reciprocating, the expected payoff to reciprocation for workers is higher than the expected payoff to non-reciprocation so the replicator function will be positive. This means that the frequency of the reciprocity trait in the worker population will be increasing, and the  $1 \geq p_w^t > p_w^1$  region will be a basin of attraction to (1,1). If  $p_w$  falls below  $p_w^1$ , however, it becomes in the Type 1 firms' best interest to choose the non-reciprocate strategy (i.e.  $Ep_f^1(n) > Ep_f^1(r)$ ). If there is a sufficient number of Type 1 firms (see Appendix) who switch to non-reciprocation at  $p_w^1$ , workers' expected payoffs

to reciprocity will fall below those to non-reciprocity (the replicator function will be negative) and reciprocity in the worker population will be declining. The level of reciprocity in the worker population will continue to decline, until  $p_w$  falls to  $p_w^{2,S}$ . Thus the  $p_w^1 \geq p_w^t > p_w^{2,S}$  region is a basin of attraction toward the equilibrium occurring at  $p_w^{2,S}$ .

At  $p_w^{2,S}$ , reciprocity in the worker population has fallen sufficiently low that it becomes in the Type 2 firms' interest to begin screening. If there is a sufficiently large number of Type 2 firms', the workers' expected payoffs to reciprocity will become higher than their expected payoffs to non-reciprocity when the Type 2 firms begin screening. Thus, the regime switch of the Type 2 firms will cause the replicator function to be positive at  $p_w^{2,S}$ . It will be in the Type 2 firms' interest to screen as long as  $p_w$  stays within  $p_w^{2,S} \geq p_w^t > p_w^{2,N}$ , and thus the level of reciprocity in the worker population will be increasing in this region, creating a basin of attraction toward the equilibrium at  $p_w^{2,S}$ . If  $p_w$  happens to fall below  $p_w^{2,N}$ , however, there will be too few reciprocating workers left in the population for it to be profitable for the Type 2 firms to continue screening. Below this point, it will be sufficiently difficult for the Type 2 firms to find a reciprocator, even with the screening technology, that the cost of screening will outweigh the expected gain to that strategy. Thus at  $p_w^{2,N}$ , Type 2 firms will switch to non-reciprocity and the workers' expected payoffs to non-reciprocity will be higher than to reciprocity and  $p_w$  will be declining in the region  $p_w^{2,N} \geq p_w^t \geq 0$ . At levels of worker reciprocity below  $p_w^{2,N}$ , both types of firms will be non-reciprocators and  $p_w$  will be declining, creating a basin of attraction toward (0,0) in the  $p_w^{2,N} \geq p_w^t \geq 0$  region.

### Stability

Note that, as in the basic model, the expanded model produces equilibria at (0,0) and (1,1) with basins of attraction toward each. It is obvious from inspection of Figure 2 that the equilibria occurring at  $p_w^1$  and  $p_w^{2,N}$  are unstable. In this simple example, the  $p_w^{2,S}$  equilibrium is a chaotic attractor. That is, if the population is at the  $p_w^{2,S}$  equilibrium and is slightly perturbed in either direction, the population will remain in the basin of attraction toward  $p_w^{2,S}$  but will repeatedly jump back and forth to positions above and below this equilibrium.

Now imagine how these dynamics would change if  $f$  had positive and continuous density *in the neighborhood* of  $p_2^r$ . This implies that rather than abrupt shifts in sign, the replicator function will transition continuously between being positive and negative. With a continuous  $p_w^{t+1}$  function, the  $p_w^1$  and  $p_w^{2,N}$  equilibria are still unstable, but the possibility emerges that the internal equilibria occurring at  $p_w^{2,S}$  will be locally stable. Recall that:

$$(16) \quad p_w^{t+1} = R_w^t + p_w^t$$

Subtracting  $p_w^{\text{int}}$  from both sides and dividing by  $(p_w^t - p_w^{\text{int}})$ , yields:

$$(17) \quad \frac{p_w^{t+1} - p_w^{\text{int}}}{p_w^t - p_w^{\text{int}}} = \frac{R_w^t}{p_w^t - p_w^{\text{int}}} + 1$$

For the population to be converging to an interior equilibrium ( $p_w^{\text{int}}$ ), the expression on the left-hand-side of Equation (17) should be a fraction—that is, the difference between the equilibrium level of reciprocity and the frequency of reciprocity in the population at time  $t+1$  and should be lower in magnitude than the difference between the equilibrium and the frequency of reciprocity at time  $t$ . This implies that, for convergence, it must be the case that:

$$(18) \quad -1 < \frac{R_w^t}{p_w^t - p_w^{\text{int}}} < 0$$

From inspection of Figure 2 one can see that, when  $p_w^{\text{int}}$  is the  $p_w^{2,S}$  equilibrium,  $(p_w^t - p_w^{\text{int}})$  and  $R_w^t$  will have opposite signs: when the population is below  $p_w^{2,S}$  (i.e.  $(p_w^t - p_w^{\text{int}}) < 0$ ), the level of worker reciprocity is increasing and thus the replicator function is positive; when the population is above  $p_w^{2,S}$ , the replicator function is negative. If the “speed of learning” parameters ( $d_1, d_2$ ) in the replicator function (Equation (3)) are chosen to be sufficiently low, it will always be the case that Equation (17) holds and thus that the  $p_w^{2,S}$  equilibrium will be locally stable.

Starting with Doeringer and Piore (1971), many observers have suggested that the way firms relate to the labor market can be categorized into two broad categories. In particular, it seems reasonable to view primary and secondary sector firms as exhibiting relatively high and low returns to reciprocity respectively (Akerlof 1982). The example used above depended on an extreme form of bi-modality, but a much less extreme bi-modal distribution would be sufficient to illustrate the possibility of the existence of a locally stable interior equilibrium.

## V. Welfare Analysis

It is easy to see that the equilibrium in which all firms and workers are involved in reciprocal employment relationships Pareto dominates the equilibrium in which no such relationships exist. In fact, it must be the case that the (1,1) equilibrium Pareto dominates all of the interior equilibria as well. This is because, when the economy is at the (1,1) equilibrium, profits and worker payoffs are maximized and costs are minimized. The (1,1) equilibrium is superior because the payoffs to mutual reciprocation are higher than those to all other types of interactions and no firm needs to pay the screening cost.

Why not, then, simply focus a welfare analysis on how economic policy could be constructed so that we could push the economy into the basin of attraction toward (1,1)? There are a few reasons why this may be an impractical policy goal. First, undoubtedly there are certain types of firms who gain little from reciprocal relations with employees.

In terms of the model presented above, these firms have very low returns to reciprocity. It is not hard to imagine, however, that there might be firms which would not want to be reciprocators, even if all of the workers were reciprocators. This can be achieved in the model by allowing the support of  $p_{ff}^{rr}$  to begin below  $p_f^{nr}$ . If we extended the range of returns to reciprocity in this manner, there would be no (1,1) equilibrium, since some firms would always want to be non-reciprocators, but the existence of a stable interior equilibrium would not be precluded. It may be unrealistic to assume that no firms will be persistent non-reciprocators, and thus a policy designed to push the economy into a perhaps non-existent (or infinitesimally small) basin of attraction toward (1,1) could be impossible to accomplish.

Another reason why it may be impractical to try to achieve universal reciprocation is that the cost of maintaining a policy which led *all* firms to be reciprocators may outweigh its benefits. Think of a policy which provided a subsidy to reciprocating firms. The subsidy would have to be attached to some identifiable proxy behavior, rather than “reciprocation” which may be hard to define or verify. For example, we might subsidize the adoption of job rotation, or an employee training program on statistical quality control. The level of the subsidy necessary to encourage the firm to become a reciprocator may be less than the gains that will be realized when the firm switches behavior. However, we cannot rule out the possibility that there will be some firms for which the necessary subsidy will be so costly that it will far outweigh the benefits to its becoming a reciprocator. If there are enough firms of the latter sort, a subsidy such as this—designed to lead all firms to becoming reciprocators—could be impractical, if not impossible, to sustain.

For these reasons, it may be more useful to consider how marginal changes in the stable interior equilibrium could be welfare enhancing. However, an analysis of the change in welfare induced by an economic policy which shifts the internal equilibrium is not as straightforward as an analysis of the movement to (1,1) from all other equilibria. This is because there are multiple externalities in the model, complicating the analysis. Since a firm’s decision to switch behavioral regimes has external effects both for the expected payoffs of workers and for the expected payoffs of other firms, the firm is not making its strategy decision on the correct boundary. If all of these externalities are unambiguously positive, then the equilibria can still be Pareto ranked by their proximity to (1,1). But if the externality has ambiguous effects, we might not be able to say whether a change in the internal equilibria is indeed welfare-enhancing.

The firm’s decision to be a reciprocator (whether or not it screens) obviously has positive external effects for workers, whose expected payoffs will go up because of the change in  $p_f$ . There is a second external effect to the firm’s decision to become a reciprocator. Because this decision will increase the number of reciprocating workers, it will also, in turn, increase the profits to all other firms (regardless of their behavior). The positive externality both to other firms and to workers that results from the firm’s decision to become a reciprocator, implies that firms are not capturing all of the surplus to their behavior.

It is the firm’s decision to begin screening which may have ambiguous welfare consequences. The transition to screening will have positive externalities for the same reason that the decision to become a reciprocator will have them—it increases  $p_f$ , thus

increasing worker payoffs, thus increasing  $p_w$ , thus increasing firm payoffs. The ambiguous external effects come from the fact that screening firms are more likely to find reciprocating workers. This means that it will be harder for the other firms, both reciprocators and non-reciprocators, to find reciprocating workers. Similarly, while it will make reciprocating workers better off (they are more likely to find reciprocators) it will make non-reciprocating workers worse off (it is harder for them to find a reciprocator). Notice, however, that the reciprocating workers are made even better off than they were when the firm decided to be a (non-screening) reciprocator, since they are more likely to be paired with the screening firm.

It turns out that when we analyze the welfare consequences of a reciprocity-enhancing policy, these positive externalities outweigh the negative effects. Consider the effects of subsidizing reciprocating firms (both those that screen and those that do not).<sup>26</sup> The increase in profit due to the subsidy would be realized by a reciprocating firm, whether or not the firm actually found a reciprocating worker. The subsidy thus increases the expected payoffs to both the reciprocation and reciprocation-with-screening strategies relative to the non-reciprocation strategy. At the internal equilibrium, firm strategies will be distributed along the returns to reciprocity as follows: firms at the low end of the  $p_{ff}^{rr}$  distribution will be non-reciprocators, an intermediate range of  $p_{ff}^{rr}$  firms will be reciprocators, and firms with the highest  $p_{ff}^{rr}$  will be screening.<sup>27</sup> Adding a subsidy to the profits of reciprocators and screeners will shift the boundaries between these three regions. Thus, with the introduction of a subsidy we will not observe any non-reciprocators becoming screeners but will, depending on the size of the shift in the boundaries and the density of firms near the boundaries, have some firms switching from non-reciprocity to reciprocity and some firms switching from reciprocity to screening.

Think first of the welfare effects these changes in firm behavior will have on workers. The introduction of the subsidy will increase the overall level of  $p_f$ , which will improve the expected payoffs to both reciprocating workers and non-reciprocating workers. But recall the potentially ambiguous welfare consequences of an increase in the number of screening firms. If a sufficiently large portion of the increase in  $p_f$  comes from an increase in  $S$ , and if the resulting decrease in the probability of a non-reciprocating worker finding a reciprocator is sufficiently high, the negative effects of the introduction of the subsidy on non-reciprocators may outweigh its positive effects on these workers. However, even if the non-reciprocating workers' loss in welfare is greater than their gain in welfare from the increase in  $p_f$  and  $S$ , the overall welfare consequences for workers will be positive. This is because, for each non-reciprocator who is displaced by this change in policy—that is, who is now less likely to realize the gains to interacting with a reciprocating firm—there is a reciprocating worker who is

---

<sup>26</sup> See the Appendix for a formal welfare analysis of this subsidy.

<sup>27</sup> There is also a special case, depending on the level of the screening cost, in which firms are only screeners or non-reciprocators at the equilibrium. Intuitively, in this case firms with low  $p_{ff}^{rr}$  are non-reciprocators and those with relatively high  $p_{ff}^{rr}$  are screeners. This two-regime equilibrium will occur when the cost of screening is sufficiently low that if it is in the firm's interest to be a reciprocator, then it will be in their interest to increase their probability of finding a reciprocator by screening.

taking his place. So if the increase in screening decreases the welfare of non-reciprocators, it must be increasing the welfare of those reciprocating workers who are now more likely to be paired with reciprocating firms. This welfare gain to reciprocators will be greater than the loss to non-reciprocators since the reciprocating workers will be gaining  $(p_w^{rr} - p_w^{mr})$  and the non-reciprocating workers will only be losing  $(p_w^{nr} - p_w^{mm})$ .

The overall effect of the subsidy for the welfare of the firms must be positive by parallel reasoning. All of the firms are made better off by the increase in  $p_w$  that results from the increase in  $p_f$ . However, non-screening firms are less likely to find reciprocators as  $S$  increases, because the screening firms are taking more reciprocating workers out of the workforce. This gain in the welfare of the screening firms must be greater in magnitude than the loss to reciprocating non-screener and to non-reciprocators because reciprocating workers are valued most highly at the screening firms (who have the highest returns to reciprocity). Therefore, even though non-screening reciprocating firms are less likely to be paired with reciprocators, the reciprocators are more likely to be in firms where they generate higher surplus. The policy will thus be unambiguously welfare-enhancing, as will all policies directed toward increasing the firms' returns to reciprocity.

## Conclusion

Although the significance of reciprocity for economics has been researched in the past, the adoption of this social norm is typically taken as exogenous. The model presented here is an attempt to formalize the notion that the firm may influence the evolution of reciprocity in society. It also provides an explanation for why we see reciprocal employment relationships developing in certain types firms and not others. Although there are multiple externalities in the model, we illustrate that a policy designed to increase the level of reciprocity at the interior equilibrium will be welfare-enhancing. In addition to formalizing the notion that cultural and economic institutions may be reciprocally influential, we hope that this model will be viewed as a step toward understanding the potential role of the work environment in shaping values, norms and personality.

## References

- Akerlof, George A. "Labor Contracts as Partial Gift Exchange," *Quarterly Journal of Economics*, November 1982, 97(4), pp. 543-569.
- Akerlof, George A. and Yellen, Janet L. "The Fair Wage-Effort Hypothesis and Unemployment," *Quarterly Journal of Economics*, May 1990, 105(2), pp. 255-283.
- Axelrod, Robert. *The Evolution of Cooperation*. New York: Basic Books, Inc., 1984.
- Baal, J. Van. *Reciprocity and the Position of Women*, Amsterdam: Van Gorcum, 1975.
- Batt, R. and E. Appelbaum. 1994. *The New American Workplace*. Ithaca, N.Y.: ILR Press.
- Berg, Joyce, Dickhaut, John and McCabe, Kevin. "Trust Reciprocity and Social History," *Games and Economic Behavior*, Vol. 10, 1995, pp. 122-142.
- Bewley, Truman F. 1999. *Why Wages Don't Fall During a Recession*. Cambridge, MA: Harvard University Press.
- Bolton, Gary and Zwick, Rami. "Anonymity versus Punishment in Ultimatum Bargaining," *Games and Economic Behavior*, 10(1), 1995, pp. 95-121.
- Bowles, Samuel and Gintis, Herbert. "The Evolution of Strong Reciprocity," Working Paper, University of Massachusetts, Amherst, 1998a.
- \_\_\_\_\_ and \_\_\_\_\_, "How Communities Govern: The Structural Basis of Pro-Social Norms," in *Economics, Values, and Organization*, Ben-Ner, Avner and Putterman, Louis, eds., Cambridge: Cambridge University Press, 1998b.
- Boyd, Robert and Richerson, Peter J. *Culture and the Evolutionary Process*. Chicago: University of Chicago Press, 1985.
- Cameron, Lisa. "Raising the Stakes in the Ultimatum Game: Experimental Evidence From Indonesia," *Economic Inquiry*, Vol. 37, January 1999, pp. 47-59.
- Cavalli-Sforza, Luigi L and Feldman, Marcus W. *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton: Princeton University Press, 1981.
- Chang, Juin-jen and Ching-chong Lai. "Carrots or Sticks? A Social Custom Viewpoint on Worker Effort," *European Journal of Political Economy*, 1999, 15, pp. 297-310.

- Davis, Douglas D. and Holt, Charles A. *Experimental Economics*. Princeton: Princeton University Press, 1993.
- Deci, Edward L. *Intrinsic Motivation*, New York: Plenum Press, 1975.
- Doeringer, Peter B. and Piore, Michael J. *Internal Labor Markets and Manpower Analysis*. Lexington, MA: Heath Lexington Books, 1971.
- Drago, Robert and Perlman, Richard. "Supervision and High Wages as Competing Incentives: A Basis for Labor Segmentation Theory," in *Microeconomic Issues in Labor Economics*, Drago, Robert and Perlman, Richard, eds. New York: Harvester Wheatsheaf, 1989.
- Eckel, Catherine C. and Grossman, Phillip J. "Altruism in Anonymous Dictator Games," *Games and Economic Behavior*, 16(2), October 1996, pp. 181-91.
- Falk, Armin and Fischbacher, Urs. "A Theory of Reciprocity," Mimeo, University of Zurich, 1998.
- \_\_\_\_\_, and Simon Gächter. "Why Firms Refuse to Hire Underbidders," *Working Paper*, University of Zurich, 1998.
- \_\_\_\_\_, \_\_\_\_\_, and Kovacs, Judit. "Intrinsic Motivation and Extrinsic Incentives in a Repeated Game with Incomplete Contracts," Working Paper, University of Zurich, 1998.
- Fehr, Ernst and Gächter, Simon. "Reciprocity and Economics: The Economic Implications of *Homo Reciprocans*," Working Paper, University of Zurich, 1997.
- Fehr, Ernst and Falk, Armin. "Wage Rigidities in a Competitive Incomplete Contract Market: An Experimental Investigation," *Journal of Political Economy*, 1999, 107, pp. 106-134.
- Forsythe, Robert, Horowitz, Joel L., Savin, N. E. and Sefton, Martin. "Fairness in Simple Bargaining Games," *Games and Economic Behavior*, Vol. 6, 1994, pp. 347-369.
- Frey, Bruno. "Shirking or Work Morale? The Impact of Regulating," *European Economic Review*, 1993, 7, pp. 1523-1532.
- \_\_\_\_\_. "How Intrinsic Motivation is Crowded Out and In," *Rationality and Society*, July 1994, 6(3), pp. 334-352.

- \_\_\_\_\_. "On the Relationship Between Intrinsic and Extrinsic Work Motivation," *International Journal of Industrial Organization*, 1997, 15, pp. 427-439.
- Güth, Werner. "On Ultimatum Bargaining Experiments—A Personal Review," *Journal of Economic Behavior and Organization*, 3(3), 1995, pp. 367-88.
- Güth, Werner, et al. "An Experimental Analysis of Ultimatum Bargaining," *Journal of Economic Behavior and Organization*, 3(4), 1982, pp. 367-88.
- Guzzo, Richard A. and Katzell, Raymond A. "Effects of Economic Incentives on Productivity: A Psychological View," in *Incentives, Cooperation and Risk Sharing*, Nalbantian, Haig R., ed. New Jersey: Rowman and Littlefield, 1987.
- Herzberg, Frederick. "One More Time: How do you Motivate Employees?" *Harvard Business Review*, 46, 53-62.
- Hoffman, E., McCabe, K., Shachat, K. and Smith, V. "Preferences, Property Rights and Anonymity in Bargaining Games," *Games and Economic Behavior*, 1994, Vol. 7, pp. 346-380.
- Ichniowski, C., Shaw, K. and G. Prennushi. "The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines," *American Economic Review*, June, 1997, 87 (3), pp. 291-313.
- Kagel, John H. and Roth, Alvin E. *Handbook of Experimental Economics*, Princeton: Princeton University Press, 1995.
- Kahneman, Daniel, Knetsch, Jack L. and Thaler, Richard. "Fairness as a Constraint on Profit Seeking: Entitlements in the Market," *American Economic Review* Vol. 76, 1986, pp. 728-41.
- Kreps, David. "Intrinsic Motivation and Extrinsic Incentives," *American Economic Review*, May 1997, 87 (2), pp. 359-364.
- Komter, Aafke E. *The Gift: An Interdisciplinary Perspective*, Amsterdam: Amsterdam University Press, 1996.
- Locke, E. A. and D. M. Schweiger, "Participation in Decision-Making; One More Look," *Research in Organizational Behavior*, 1 (10), pp. 265-339.
- Mauss, Marcel. *The Gift: Forms and Functions of Exchange in Archaic Societies*, New York: W. W. Norton and Co., 1967.
- Miller, K. and P. Monge. "Participation, Satisfaction and Productivity: A Meta-Analytic Review," *Academy of Management Journal*, 29, pp. 727-53.

Schrift, Alan D. *The Logic of the Gift: Toward an Ethnic Generosity*, New York: Routledge, 1997.

Thaler, Richard H. "Anomalies: The Ultimatum Game," *Journal of Economic Perspectives* Fall, 1988, 2(4), pp. 195-206.