Human Relations in the Workplace

Julio J. Rotemberg

Massachusetts Institute of Technology

This paper seeks to understand what motivates workers to be altruistic toward one another and studies whether firms benefit from encouraging these "human relations" in the workplace. The paper first proposes that feelings of altruism can be individually rational in certain settings in which the variables controlled by the workers are strategically linked. The paper then studies what this implies for equilibrium altruism in two situations. The first has workers who are paid as a function of joint output. The second is the relationship between subordinates and their supervisors.

People have feelings for those they work with. This raises two questions: first, whether these feelings affect performance on the job and, second, what gives rise to these feelings. In this paper I consider both the effects and the causes of one particular set of feelings. In particular, I focus on the happiness people experience when good things happen to their coworkers. I thus analyze the types of interactions at work that lead to altruistic feelings toward fellow employees. I also study whether firms can benefit from encouraging these feelings.

The idea that personal feelings (or "human relations") affect performance in the workplace is an old one, dating back at least to Mayo (1933). The famous "Hawthorne experiments" reported by Roethlisberger and Dickson (1939) appeared to confirm the importance of these personal relations in industrial settings. Subsequently, a large number of empirical studies were conducted relating productivity in

I wish to thank Deborah Ancona for helpful conversations; Gary Becker, Drew Fudenberg, David Laibson, Casey Mulligan, Bob Pindyck, Rob Porter, Matthew Rabin, Garth Saloner, Jeremy Stein, Jean Tirole, and two anonymous referees for comments; and the National Science Foundation for research support.

[*Journal of Political Economy*, 1994, vol. 102, no. 4]

small groups to the feelings reported by the members of the group. One particular feature of groups that has been explored in several studies (see Goodman, Ravlin, and Schminke [1987] for a discussion) is group cohesiveness, a term whose definition varies but that often means the extent to which members of the group express that they like other members of the group. Unfortunately, these studies of the relation between cohesiveness and productivity have delivered mixed results.

Attraction between people (or liking of each other) need not be identical to altruism. However, the two feelings appear linked in practice. One reason is that once a person is altruistic toward another, he or she will generally enjoy the other's company. This personal interaction gives the altruist an opportunity to be nice and thus to gain from the happiness thus generated. Another reason for attraction to be related to altruism is that, as I shall clarify further below, it may well be beneficial to oneself to become altruistic toward those one finds oneself attracted to.

Finally, I shall argue that some aspects of the empirical literature on cohesiveness are easy to understand if one equates people's expression of liking for each other (i.e., those that lead groups to be classified as highly cohesive) with feelings of altruism. In particular, it becomes straightforward to explain why the relationship between cohesiveness and productivity is ambiguous. Altruism toward fellow employees may lead an employee to work little or to work hard. It has the former effect if a cutback in an individual's effort allows fellow employees to reduce their own effort while keeping their income constant. It has the latter effect if, by working harder, an employee raises the income of fellow employees.

This ambiguity is the same as that analyzed by Holmström and Milgrom (1990) in their study of "collusion" in organizations. They focused on the effects on firm profitability of allowing workers to collude, that is, to write binding contracts with each other. As with altruism, these binding contracts also lead workers to take actions that enhance each other's welfare. Holmström and Milgrom show that when effort would be enhanced by competition, as in Lazear and Rosen (1981), collusion is bad for the firm because it leads both employees to exert lower effort. On the other hand, if the firm can observe only collective output so that compensation is tied to group performance, collusion is beneficial to the firm.

One issue that arises at this point is why one should study altruism in the workplace if its effects are similar to those of contracts. One reason is that contracts between workers, particularly if they hurt the employer, are not enforceable by the courts. However, the contracts might be self-enforcing as a result of the repeated nature of the interaction. One key difference between this self-enforcement and altruism is that altruism works even in the last period of the relationship. Thus altruism can explain why workers take actions whose principal purpose is to benefit their coworkers even on their last day at work.

If altruism does indeed affect performance within the workplace, it becomes important to ask what brings it about. In this paper, I develop a model in which people become altruistic when doing so is in their self-interest. Formally, I treat the degree of altruism as a choice variable: an individual starts with a utility defined over his own welfare and chooses whether to change his utility function and become altruistic.[1] The object of this formal model is to develop a theory of the conditions under which altruism will arise in work settings.

The assumption that people choose their altruism in their own self-interest, which may be controversial, can be thought of in two additional ways. First, one can think of individuals as taking actions (such as inviting others to dinner) that modify their attitude toward others. Viewed in this way, the model is akin to the rational addiction model of Becker and Murphy (1988), where individuals know that the initial consumption of an addictive substance will change their future attitude toward the substance. Second, as I elaborate further below, natural selection might favor genes that lead people to become altruistic when doing so benefits them.

I now turn to the benefits that the individual derives from his altruism. Without repeated game considerations, from which I abstract, altruism is not in the individual's self-interest unless it can be credibly demonstrated. The reason is that altruism leads the individual to take actions that maximize a combination of his own and the other's utility function, and this results in a lower level of own utility than if the individual had been purely selfish.

On the other hand, consider the case in which A can prove to B

[1] The sort of opportunistic change in tastes I consider is also featured in Raub (1990). Other related papers include Akerlof (1983), which shows how people benefit from becoming loyal, and Frank (1987). Frank shows that people capable of guilt will try to play Prisoner's Dilemma games with others who are similarly capable since the fear of guilt feelings will induce both to cooperate. A similar idea is present in the book by Gauthier (1986), who assumes that people can choose to become "constrained maximizers." If they do so, they cooperate in Prisoner's Dilemma games if they think that their partner is also a constrained maximizer. Otherwise they defect. Coleman's (1990) self-interested change in tastes is somewhat different since it involves an actor who becomes happy by "changing himself to be satisfied with the world" (p. 517). Coleman also considers actors who relinquish control of their actions to groups of individuals. This is more similar to what is considered here since these groups have objectives potentially different from those of the narrowly conceived individual. For a model in which altruism is a choice variable, though not with the purpose of increasing material payoffs, see Mulligan (1993).

that he feels altruistic toward him. Knowing this, B knows that A's behavior will be different. This, in turn, will sometimes affect B's behavior. If B is led to modify his behavior in a way that benefits A, the initial decision by A to feel altruistic toward B will have been a smart one. I shall show that the condition under which this occurs is related to the strategic complementarity of Fudenberg and Tirole (1984) and Bulow, Geanakoplos, and Klemperer (1985). This condition requires that, if each individual takes one action and these actions are normalized so that increases in one person's action raise the other person's welfare, a marginal increase in A's action leads to an increase in B's.

In this case, it is B's belief that A will take subsequent actions that differ from those of a selfish A that make altruism beneficial for A. This belief makes sense only if altruism lasts from the moment it is detected until the subsequent actions are taken. It must thus be irreversible, at least to some degree.[2] In practice, one does see people who stop liking one another, and as Mulligan (1993) points out, altruism has the potential to turn into its opposite, the desire to see the other person suffer losses in utility. These further changes in emotions seem to arise as conditions within the relationship change, usually in ex ante unpredictable ways. To keep the analysis simple, my model has individuals interacting only once so that these later changes in emotions play no role.

This basic idea that altruism can be privately beneficial is closely related to Becker (1974). Becker's "rotten kid theorem" establishes that the interaction between a selfish and an altruistic individual can lead to Pareto-optimal allocations. In particular, the selfish individual will, under certain circumstances, take actions that maximize joint income. This result hinges on three assumptions. First, the altruistic individual must be so rich and so altruistic that, in equilibrium, he makes positive transfers to the selfish one. Second, the altruistic individual must act *after* the selfish one. These assumptions imply that the altruist will lower his transfers to the selfish individual when the latter reduces joint income.[3] Finally, the selfish individual's utility must be linear in the ex post transfers by the altruistic one (Bergstrom 1989).

The rotten kid theorem does not establish that the altruist gains from his altruism. In a closely related paper, Becker (1976) argues that altruistic individuals can, indeed, end up better off than selfish ones. His argument is largely informal, though it too is based on an example in which the altruist would, ex post, punish the selfish

[2] It is thus a form of commitment.

[3] Further discussion of these two assumptions can be found in Hirshleifer (1977).

individual if the latter failed to take actions that benefit the altruist. On the other hand, Bernheim and Stark (1988) provide an example in which selfish individuals do better than altruists.

Given these contradictory findings, I devote Sections I and IV to clarifying when altruism is individually rational.[4] To avoid some of the asymmetries introduced when players take actions sequentially, I focus mostly on games with simultaneous moves. Only in Section IV do I discuss sequential moves. With this change in the timing of moves, strategic complementarity ceases to be sufficient for equilibrium altruism, although it remains a force that promotes such altruism.

After considering rational altruism within abstract simultaneous move games in Section I, I turn to its applications in the workplace. This boils down to an analysis of whether the strategic variables of different employees tend to be strategic complements or substitutes in different settings. When they are complements, altruism should arise endogenously.

The first relationship I consider is that between employees who work as a team in that their payment depends only on their joint output. I show that, as a result, altruism tends to arise and raise productivity. From the firm's perspective this is a good thing. This leads to the question whether the firm can foster even more altruism by making it possible to socialize on the job. I consider this issue explicitly for two reasons. First, it allows me to explain the results of the Hawthorne experiments reported by Roethlisberger and Dickson (1939). They suggest that changes in incentive payments together with the creation of an atmosphere conducive to friendship helped productivity more than either change on its own. Second, discussion of the relationship between socializing and altruism allows me to explain why attraction tends to lead to altruism. Section II also contains an extended discussion of some additional evidence on the relationship between cohesiveness and productivity.

In Section III, I consider relations of authority. Several authors including Homans (1950) have pointed out that relations of authority tend to be less friendly than relations among more equal coworkers. I show that, in the case of supervisors whose main role is to monitor their subordinates' actions for purposes of compensation, one should expect altruism to be absent. I also consider an authority relationship described in Crozier (1964) in which a "team leader" has no authority to set compensation but does have some control over the pace of work. I focus on this relationship because it has an unexpected fea-

[4] Altruism is rational in my model in the same sense that addiction is rational in Becker and Murphy (1988).

ture. The workers in this setting generally reported that they did not like each other (which I equate with lack of altruism). The exception was that team leaders liked their coworkers. My model can explain this asymmetry on the basis of the benefits to the team leader from gaining the trust of her coworkers. Section IV considers sequential games, and Section V offers some conclusions.

I. Rational Altruism

In this section, I show how individuals can benefit by changing their tastes toward altruism. I first take up the case of two individuals, each of whom must make a continuous choice. I then deal briefly with the Prisoner's Dilemma problem and the case of n individuals.

A. *Continuous Choices of Two Individuals*

The two individuals, A and B, carry out actions a and b, respectively. The initial utility functions of A and B are, respectively, $\alpha(a, b)$ and $\beta(a, b)$. These utility functions can be thought of as being the "true" welfare (or, in the terminology of Rabin [1993], "material payoffs") of A and B so that α depends only on what happens to A individually, and similarly for β. Thus in my simple settings, α depends on A's income and on his own effort but not on B's income or effort.

If the utility functions remain α and β, the Nash equilibrium pair of actions a^n and b^n solve the following standard equations:

$$\alpha_a(a^n, b^n) = 0,$$

$$\beta_b(a^n, b^n) = 0,$$

where subscripts denote partial derivatives. The Nash equilibrium is generally not Pareto optimal. Pareto optimality requires that, for some positive λ, the pair of actions (a, b) satisfy

$$\begin{aligned} \alpha_a(a, b) + \lambda\beta_a(a, b) &= 0, \\ \alpha_b(a, b) + \lambda\beta_b(a, b) &= 0. \end{aligned} \tag{1}$$

Thus Pareto optimality holds at the Nash equilibrium only if $\alpha_b(a^n, b^n) = 0$ and $\beta_a(a^n, b^n) = 0$. I shall be concerned with cases in which Pareto optimality does not hold so that these derivatives are nonzero at the selfish Nash equilibrium.

I now turn to the case in which B feels altruism toward A. This means that B's psychological well-being depends on A's material payoff α. The standard formulation in the case in which altruism is exogenous thus makes B's utility be a linear combination of β and α with a higher weight on α signifying a higher degree of altruism. In the second stage of my model, altruism has already been chosen, and I follow this standard specification. Thus at this point individual actions maximize utility functions given by

$$U_A = \alpha(a, b) + \gamma_A \beta(a, b), \tag{2}$$

$$U_B = \beta(a, b) + \gamma_B \alpha(a, b). \tag{3}$$

I also assume that there is a first stage in which altruism is chosen, and this raises an issue. Suppose that B picks γ_B to maximize either U_B or a more general function such as

$$U_B' = \beta + \gamma_B(\alpha - \alpha^o) - f(\gamma_B), \tag{4}$$

where $f(\gamma_B)$ captures the direct costs of changing altruism discussed by Montgomery (1991) and Mulligan (1993).[5] Then changes in γ_B generally change the level of utility even when altruism has no behavioral consequences.[6] In many applications, particularly those concerning the family, these direct effects may well constitute the principal reason for an agent to change his altruism. Here, however, I wish to abstract from these direct effects.

I thus assume that B chooses his altruism to maximize his material payoff β itself.[7] After having chosen to be altruistic, he does behave as though his utility is given by (3) so that the altruism is genuine. In this specification, the individual uses a selfish utility function to evaluate the desirability of altruism. There are a number of interpretations for this formulation. As in Coleman (1990), one can think of an "inner" self that is selfish and relinquishes control of actions to an "outer" self. What the inner self can do, however, is to mold the preferences that guide the outer self's actions. Thus the inner self can make the outer self altruistic, and this altruism becomes genuine because the inner self cannot change the outer self's preferences too

[5] Mulligan's discussion of costs is particularly germane to the issues considered here. Following Smith (1853), he stresses the cost of imagining how the other person will feel. The difficulties involved could explain why altruists within the workplace do favors only within a narrow range. The altruist senses only the effect of these favors and not that of others.

[6] These direct effects can be dramatic. If B were to maximize U_B, he would choose to set γ equal to infinity if α were positive and equal to negative infinity if α were negative and thereby get unbounded utility. Related problems with having people choose altruism parameters of this form are discussed in Bernheim and Stark (1988).

[7] This specification has the disadvantage that it makes it hard to interpret the act of becoming altruistic along the lines of Becker and Murphy (1988). Individual B cannot easily be thought of as choosing an action that modifies his utility function today in the same manner that it modifies his utility tomorrow. For this purpose, it would be better to let B choose his altruism parameter γ_B to maximize (4). For this reason, I also briefly considered this alternative in a previous version of this paper. I show there that, even when people are maximizing functions such as (4), they are more likely to become altruistic in equilibrium if β itself rises as a result.

rapidly. A second interpretation is evolutionary. If emotional reactions are guided by genes, natural selection might favor the reproduction of individuals whose emotions change in self-interested ways.[8] A related possibility is that natural selection favors those genes that lead people to imitate the behavior of individuals who appear successful. Insofar as people appear successful when their material payoffs are high and it is possible to infer the γ's of successful individuals from their behavior, this will lead people to choose γ's in a way that maximizes material payoffs.

After the γ's are chosen, A and B choose a and b to maximize (2)–(3). The Nash equilibrium actions thus satisfy

$$\alpha_a + \gamma_A \beta_a = 0, \tag{5}$$

$$\beta_b + \gamma_B \alpha_b = 0. \tag{6}$$

If one compares (1), (5), and (6), it is apparent that, in the limit in which γ_A and γ_B equal one, the Nash equilibrium with altruism is Pareto optimal. Suppose in addition that the solution to (5) and (6) when $\gamma_A = \gamma_B$ is monotone in the common value of γ. It then follows that increasing γ leads to actions that are closer to the Pareto optimum. This explains why simultaneous increases in altruism are beneficial to the individuals. However, I am more concerned with whether an individual would unilaterally decide to become altruistic.

To study this, I now analyze the effect of changes in γ_B on the actions a and b. Consider first the case in which A is unaware of B's change of heart. Then a remains constant and the change in b can be obtained by differentiating (6):

$$(\beta_{bb} + \gamma_B \alpha_{bb})db = -\alpha_b d\gamma_B. \tag{7}$$

The second-order conditions for utility maximization imply that β must decline (to second order) when b is changed in this way. So B is worse off when we use this measure of his ex ante welfare. By contrast, the increase in γ_B makes A better off since

$$\alpha_b db = -\frac{(\alpha_b)^2}{\beta_{bb} + \gamma_B \alpha_{bb}} d\gamma_B > 0.$$

[8] This begs the question of why natural selection does not favor genes that lead individuals to act as though they were altruistic without actually experiencing any affection or empathy. Perhaps, as suggested by Akerlof (1983) about loyalty, it is somehow more difficult to pretend altruism than it is to experience it. Being altruistic might make it possible to carry out acts that give much utility at low cost to oneself by allowing one to "put oneself in the other's shoes." But having the ability to do this may then make it difficult to ignore any negative effect one's subsequent acts have on the other.

I now turn to the case in which B can demonstrate his altruism. Given that small increases in γ_B from zero make B worse off, any first-order gains to B must be due to the changes this altruism induces in A's behavior. But A will change his actions only if B can demonstrate his unselfish regard credibly, so that A is sure that B's actions will be affected. Such credibility requires that it be difficult (if not impossible) for a person in B's position who does not feel altruistic toward A to mimic the signal emitted by a true altruist.

What signals should A believe? There are at least two types of somewhat credible signals. The first type, body language, is stressed by Homans (1950), who says that "In deciding what sentiments a person is feeling, we take notice of slight, evanescent tones of voice, expressions of his face, movements of his hands, ways of carrying his body, and we notice these things as part of a whole. . . . From these wholes we infer the existence of internal states of the human body and call them anger, irritation, sympathy. . . . [Yet] we act on our inferences, on our diagnoses of the sentiments of other people, and we do not always act ineffectively" (p. 39). Individual B can also signal his altruism through favors and presents. This idea is pursued by Camerer (1988), who gives several reasons why gifts can be credible signals of altruism. One of these reasons is that it may be more difficult for a person who does not actually care for another to deliver equally touching favors and gifts.

For simplicity, I neglect the costs of signaling altruism. The effect of a change in γ_B on both actions can then be obtained by differentiating (5) and (6):

$$\mathbf{D}\begin{pmatrix} da \\ db \end{pmatrix} = \begin{pmatrix} 0 \\ -\alpha_b \end{pmatrix} d\gamma_B,$$

where

$$\mathbf{D} = \begin{pmatrix} \alpha_{aa} + \gamma_A \beta_{aa} & \alpha_{ab} + \gamma_A \beta_{ab} \\ \beta_{ab} + \gamma_B \alpha_{ab} & \beta_{bb} + \gamma_B \alpha_{bb} \end{pmatrix}.$$

Thus the effect on a and b is given by

$$da = \frac{\alpha_b(\alpha_{ab} + \gamma_A \beta_{ab}) d\gamma_B}{|\mathbf{D}|} \tag{8}$$

and

$$db = \frac{-\alpha_b(\alpha_{aa} + \gamma_A \beta_{aa}) d\gamma_B}{|\mathbf{D}|}. \tag{9}$$

If we use the equilibrium conditions (5) and (6) to substitute for α_a and β_b, the effects on α and β are given by

$$\begin{aligned}\frac{d\alpha}{d\gamma_B} &= \alpha_a\frac{da}{d\gamma_B} + \alpha_b\frac{db}{d\gamma_B}\\ &= \frac{-\gamma_A\beta_a\alpha_b(\alpha_{ab} + \gamma_A\beta_{ab}) - \alpha_b^2(\alpha_{aa} + \gamma_A\beta_{aa})}{|\mathbf{D}|}\end{aligned} \tag{10}$$

and

$$\begin{aligned}\frac{d\beta}{d\gamma_B} &= \beta_a\frac{da}{d\gamma_B} + \beta_b\frac{db}{d\gamma_B}\\ &= \frac{\alpha_b\beta_a(\alpha_{ab} + \gamma_A\beta_{ab}) + \gamma_B\alpha_b^2(\alpha_{aa} + \gamma_A\beta_{aa})}{|\mathbf{D}|}.\end{aligned} \tag{11}$$

To obtain a determinate sign for $|\mathbf{D}|$, I suppose that the solution is stable under the usual tatônnement process, where da/dt depends positively on dU_A/da and db/dt depends positively on dU_B/db. This stability, which can be viewed as implying that people converge to the Nash equilibrium when they act in a fairly myopic manner, implies that the matrix $\mathbf{D}$ is negative semidefinite.[9] So, as long as $\mathbf{D}$ is not singular, its determinant is positive.

I now consider the effect of changes in γ_B on ex ante welfare. As long as γ_A is zero, A benefits from B's increased altruism since the sign of $d\alpha/d\gamma_B$ is the sign of $-\alpha_{aa}\alpha_b^2$, which is positive. The sign of $d\beta$ when the two γ's are zero is equal to the sign of $\alpha_b\beta_a\alpha_{ab}$. One simple way of interpreting this sign is to normalize the actions so that α_b and β_a are positive at the Nash equilibrium. So, at the Nash equilibrium, each agent wants the other to raise the level of his action. Given this normalization,[10] $d\beta/d\gamma_B$ has the sign of α_{ab}. This second derivative is positive if a selfish A would want to increase a in response to an increase in b. In Bulow et al.'s (1985) terminology, a positive α_{ab} makes a and b "strategic complements" from A's point of view.

[9] The reason is that, locally, the system of differential equations governing a and b is then given by

$$\begin{pmatrix} da/dt \\ db/dt \end{pmatrix} = \mathbf{D}\begin{pmatrix} a - a^* \\ b - b^* \end{pmatrix},$$

where a^* and b^* are the equilibrium values.

[10] In Bulow et al.'s (1985) work on strategic interactions between firms, there is no natural analogue to this normalization. For this reason, their results depend not only on whether actions are strategic complements or not but also on whether prices or quantities are the strategic variable, whether there are economies or diseconomies across markets, etc.

This strategic complementarity then leads B to become altruistic toward A.

The intuition for this result is straightforward. If B starts to feel altruistic toward A, he will move b in the direction A likes. Normalizing the actions so that this direction is also the direction in which A must move his action to benefit B, we see that, if the actions are strategic complements, A does indeed move his action in this direction and B benefits from a positive γ_B. If, instead, A views the actions as strategic substitutes, B loses when he lets A know that he cares for him because this leads A to adjust a in the direction that B finds deleterious.[11] So, with these preferences and strategic complementarity, purely selfish equilibria do not exist; individuals would deviate from pure selfishness and become somewhat altruistic.

Equilibrium preferences can be derived as follows. In a Nash equilibrium, each individual picks his own γ to maximize his own material payoff, taking as given his conjecture of the other individual's choice of γ. Since these conjectures must be correct in equilibrium, each individual's γ is optimal given the γ picked by the other agent. I now show that equations (10) and (11) imply that, when the actions are strategic complements, there generally exists an equilibrium in which both γ's take values between zero and one. As we saw, these equations imply that the best response to a γ of zero is a strictly positive γ. Because strategic complementarity implies that β_{ab} is positive as well, B benefits from raising his γ above zero also in the case in which γ_A is itself positive.

Under plausible conditions, no individual would ever find it optimal to pick a γ larger than one. Suppose, for instance, that both individuals have the same constant marginal utility of income. Then a γ_B greater than one would lead B to give all his income to A, and this would not be good for B's ex ante welfare. The result is that the best response to any γ between zero and one is a γ between zero and one. Because these best responses lie in a compact set, the Kakutani fixed-point theorem establishes that there exists an equilibrium in which both γ's take values between zero and one.

One can establish the existence of an equilibrium with γ's between zero and one even more directly if one assumes that α and β are concave symmetric functions so that $\alpha(x, y) = \beta(y, x)$. To see this, consider $d\beta/d\gamma_B$ at the point at which both γ_A and γ_B equal one. Because at this symmetric point a and b are equal (so that α_b and β_a

[11] Bernheim and Stark (1988) provide an example in which increased altruism by B makes B worse off for this reason. Their example is based on the "Samaritan's Dilemma": B's altruism leads A to expect larger transfers in the future. This expectation in turn induces A to consume more in the present period (which is inefficient).

are equal), (11) implies that the sign of $d\beta/d\gamma_B$ at this point is equal to the sign of $2\alpha_{ab} + \alpha_{aa} + \alpha_{bb}$. The concavity of α guarantees that this is negative so that B would choose a γ smaller than one if γ_A were equal to one. This means that the best response to a γ of one is below the 45-degree line. Since the best response to a γ of zero is above this line, the intermediate value theorem establishes that there exists an intermediate γ for which the best response is equal to γ. This is a symmetric Nash equilibrium.

Figure 1 depicts the best response function of B, denoted $\gamma_B(\gamma_A)$, as well as the 45-degree line. If α and β are symmetric, equilibria can be found at the intersection of these two lines. There is, unfortunately, no guarantee that there will be only one such intersection. What can be seen from the figure is that the first intersection has some desirable properties. First, as the arrows demonstrate, it is stable under the usual tatônnement process. Second, the equilibrium γ at this equilibrium rises if the best response function rises. In other words, changes in α and β that raise $\gamma_B(\gamma_A)$ to $\gamma_B'(\gamma_A)$ and thus raise each individual's altruism for a given level of partner altruism lead to more altruism at this equilibrium.

B. *The Discrete Prisoner's Dilemma*

In this subsection I consider the standard model of conflict, namely the Prisoner's Dilemma.[12] The payoffs in this game are given by

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{bmatrix} 1, \quad 1 & -l, 1+g \\ 1+g, \; -l & 0, \quad 0 \end{bmatrix}, \end{array} \tag{12}$$

where l and g are strictly positive.[13] In the usual definition, the total payoff from cooperation, two, exceeds that from defection, $1 + g - l$.

Since each player's optimal action (D) does not depend on the other player's action, there appear to exist no strategic complementarities in this game. But endogenous altruism can still arise because, depending on the parameters, a weaker form of strategic complementarity may be present. In particular, the benefit from playing D instead of C generally depends on whether the other player plays D or

[12] The analysis of this game is related to Raub (1990). The difference is that I provide conditions on the parameters of a Prisoner's Dilemma game that imply that cooperation arises endogenously whereas Raub does not.

[13] The zeros and ones are simply normalizations. In any symmetric Prisoner's Dilemma, one can always subtract the same amount from each entry to make the $\{D, D\}$ entry equal to $\{0, 0\}$ and then multiply all the entries by a positive constant to make the $\{C, C\}$ entry equal to $\{1, 1\}$. The resulting game is strategically equivalent.

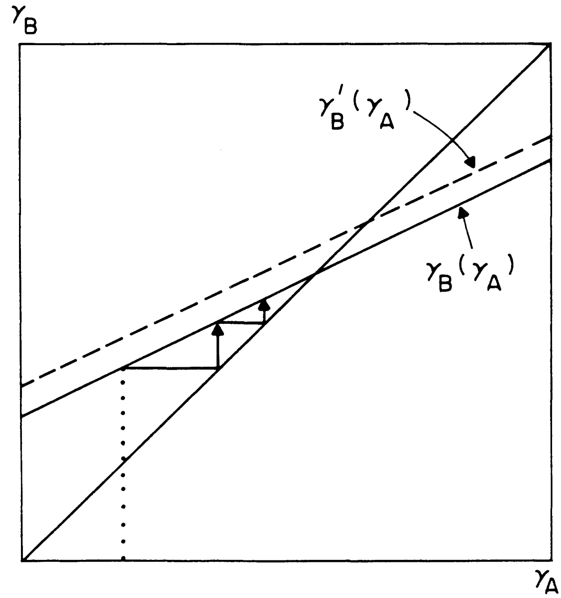

FIG. 1.—Symmetric equilibrium with two players

C. When A plays D, B's benefit from playing D instead of C equals l. When A plays C, that benefit equals g. So if l exceeds g, there exists a weak form of strategic complementarity in that each player's benefit from defecting to D is larger when the other player also plays D.

This turns out to induce actual strategic complementarity once the players have a positive level of altruism. To see this I use the material payoffs in (12) to compute the U's as in (2) and (3). They equal

$$\begin{array}{c c} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{bmatrix} 1+\gamma_A, \quad 1+\gamma_B & -l+\gamma_A(1+g),\ 1+g-\gamma_B \\ 1+g-\gamma_A l,\ -l+\gamma_B(1+g) & 0, \quad 0 \end{bmatrix} \end{array}. \tag{13}$$

If l exceeds g, there exists a range of γ's between $g/(1 + l)$ and $l/(1 + g)$. Suppose that A, the row player, has chosen γ_A in that range. Given this γ, A's best response to C is C (because $1 + \gamma_A$ is greater than $1 + g - \gamma_A l$) and his best response to D is D (because $\gamma_A(1 + g) - l$ is negative). Thus, given this γ_A, A's actions really are strategically complementary to B's. As I show in the Appendix, the result is that B and A will find it in their best interest to become altruistic.

C. *Symmetric* n-*Player Games*

In this subsection I consider briefly the case in which n players each take an action a^j. I assume that the material payoffs for player j are given by $f^j(a^1, \ldots, a^j, \ldots, a^n)$. The functions f^j are symmetric in that they are invariant with respect to permutations concerning players other than j and in that one can obtain f^i from f^j by simply permuting i and j's actions. I let player j choose the $n - 1$ values of γ_i^j that denote the altruism that j feels for i. In particular, after his choice of γ's, j maximizes

$$\sum_{i=1}^{n} \gamma_i^j f^i(a^1, \ldots, a^j, \ldots, a^n),$$

where γ_j^j cannot be chosen and is equal to one.

After the altruism parameters have been chosen, the first-order condition for player j's choice of action is

$$\sum_{i=1}^{n} \gamma_i^j f_j^i = 0. \tag{14}$$

Suppose that all players except for j take the same action, and denote this action by a^{-j}. Then the symmetry of the f's implies that (14) can be written as

$$f_j^j(a^{-j}, a^j) + f_j^{-j}(a^{-j}, a^j) \sum_{i \neq j} \gamma_i^j = 0, \tag{15}$$

where f^{-j} represents the material payoffs of any player except j. Thus j's optimal action depends only on the sum of his γ's. Now consider a symmetric situation in which $\Sigma_i \, \gamma_i^j$ is the same for all j. The action that constitutes a symmetric Nash equilibrium must then ensure that the a^j that solves (15) is equal to a^{-j}. This is indeed a Nash equilibrium since this action satisfies the first-order condition (14) for all players.

What is interesting about this equilibrium is that it does not change if one changes the γ's of an individual while keeping the sum of his γ's constant. One can achieve the same symmetric outcome whether all players have the same altruism toward all others or whether each feels altruistic toward only one individual. This result obviously hinges on the inability of each player to modify his actions so that they benefit only one player as opposed to benefiting all players symmetrically.

Now consider the stage in which the individuals choose their altruism parameters to maximize their own f's. Consider again the normalization such that increases in a^j make other players better off. The actions are then strategically complementary if an increase in a^j leads other players to raise their own a. With this strategic complementarity, altruism will once again be strictly positive at a Nash equilibrium. The reason is that players would deviate from an outcome in which all γ's are zero. A player, say player j, would raise some of his own γ's because that would credibly lead him to raise his own a and would thereby convince others to raise theirs. The same logic that leads to equilibria with altruism in the case of two players extends to this n-player case. For this reason, the examples I consider below deal only with two players.

What remains indeterminate, at least in the case of symmetric games, is the composition of j's γ's since the other players react only to the sum of the γ_i^j. To determine this composition, one must change this model somewhat. Following Mulligan (1993), one could assume that raising each individual γ is costly. If these costs are convex, the least-cost method for ensuring a certain sum of γ's is to make them all equal.

So far, I have considered only relatively abstract games. In the next sections, I study specific features of the payoffs that arise inside firms. These payoffs are made to depend on the relationship of the people within the organization. Before I develop the applications inside firms, it is important to stress that personality, which I have neglected, probably also plays a role in the determination of the equilibrium values of α and β. In particular, the private costs of the actions a and b are likely to depend on the individual. For instance, an individual who is not very interested in career advancement (perhaps because he finds his leisure very valuable) would have a different attitude toward changing his effort than one who is very keen on professional progress. Thus one should not expect the pattern of friendship and altruism to depend solely on the nature of the interactions in the workplace. With this caveat, I keep personality factors constant and analyze the effect of these interactions on altruism.

II. Team Members

Often, it is not possible to disentangle the individual contributions of employees who produce something in common. As Holmström (1982) shows, if they are paid as a function of total output alone and they are selfish, their effort is suboptimally low. The reason is that each employee neglects the positive effects on the other employee's compensation of an increase in his own effort. In subsection *A*, I show that altruism will often emerge endogenously in this setting.

Members of a productive team may or may not have the opportunity to socialize at work. In subsection *B*, I show the effect on altruism of having opportunities to socialize. This subsection also demonstrates that situations in which one individual is attracted to another tend to lead to endogenous altruism. It thereby establishes a link between altruism and the cohesion of working teams. Subsections *C* and D consider experimental evidence on the connection between cohesiveness and productivity.

A. *The Effects of Joint Production*

Suppose that output is given by the symmetric, monotone, and increasing function $f(a, b)$, where a and b once again represent the two actions. If, together, the two employees receive a price of one for their output and they split evenly the resulting revenue, the individual payoffs are

$$\alpha(a, b) = \frac{f(a, b)}{2} - e(a), \qquad \beta(a, b) = \frac{f(a, b)}{2} - e(b), \tag{16}$$

where the e functions capture the cost of effort with $e' > 0$ and $e'' > 0$ (primes denote derivatives). The selfish Nash equilibrium then has $f_a = f_b = 2e'$, whereas the optimum requires that f_a and f_b be equal to e' itself. Since the marginal product of effort is smaller at the joint optimum, the Nash equilibrium must have a smaller level of effort.

If the two employees act as though they had preferences given by (2) and (3), the equilibrium effort satisfies instead

$$f_a = \frac{2}{1 + \gamma_A} e', \qquad f_b = \frac{2}{1 + \gamma_B} e',$$

so that the marginal product of each type of effort declines as the γ's rise. In the limit in which the two γ's equal one, we obtain the same outcome as at the optimum.

Since α_{ab} equals f_{ab}, we can be sure that there is some altruism in equilibrium if, as in the Cobb-Douglas case, f_{ab} is positive. The previous analysis also establishes that the two γ's are less than one at the Nash equilibrium, where each agent chooses his level of altruism taking his conjecture of the other's choice of γ as given.

Because equilibrium altruism raises output and traditionally measured productivity, it tends to be in the firm's best interest as well. The firm benefits if the price it pays the team per unit of output (which I normalized to one) is less than the value to the firm of an addition to output. Such a difference can be expected to arise as long as the production of output involves some proprietary knowledge or even the use of some of the firm's capital, so that the firm is generally better off if the employees behave altruistically. In the next subsection, I explore a mechanism the firm may have available that increases altruism among coworkers.

B. *Socializing on the Job*

A person may like the social company of another without necessarily feeling altruistic toward him. Liking somebody's social company means only that one derives utility from the other's presence, not that one's utility is increasing in the other's utility. Yet there appears to be an empirical connection among attraction, sociability, and altruism. This subsection shows how the opportunity to socialize can give rise to endogenous altruism.

Suppose that A likes the social company of B. He gets z extra units of utility if B spends time socializing with him. To obtain endogenous altruism, I assume that, while socializing with B, A has the opportunity to raise B's happiness at some cost to himself. For example, A might accede to B's wishes while socializing, or he may spend resources more directly. These favors to B cost i to A, but A incurs these costs only when he actually socializes with B.[14] I suppose that B has a strictly positive cost c of socializing with A if A has not spent resources. If A does spend i, B's cost is only $c - f$.

I suppose that B gets to decide whether to socialize with A or not but that he does not know at this point whether or not A will spend i to improve the interaction. I treat the two decisions—B's decision to socialize, S, and A's willingness to spend i if they actually socialize, P—as simultaneous.[15] If we denote the null decisions for each player by N, the payoffs are

$$\begin{array}{cc} & \begin{array}{cc} N & \quad S \end{array} \\ \begin{array}{c} N \\ P \end{array} & \begin{bmatrix} 0, 0 & z, \quad -c \\ 0, 0 & z - i, f - c \end{bmatrix}. \end{array}$$

Without altruism, N is a dominant strategy for A so that both play N at the Nash equilibrium. For B, socializing becomes more attractive when A plays P as long as f is positive. Thus a positive f ensures that B's actions are strategic complements vis-à-vis A's. However, a positive f is not sufficient for altruism to emerge in equilibrium. Because the game is discrete, this strategic complementarity has an effect only if it is quantitatively important. In particular, altruism arises if and only

[14] Essentially identical results obtain if the cost is incurred whether A and B socialize or not.

[15] Given that A spends resources only once the interaction takes place, it might seem plausible to let A choose after B. This has no effect on the analysis.

if f is larger than c. With $f - c$ and $z - i$ positive, A gains by setting γ_A slightly greater than i/f. This ensures that A plays P, which leads B to socialize and thereby makes A better off. Thus lowering the cost of socializing c to the point at which $f - c$ is positive may be beneficial to the firm.

While this shows that A benefits from becoming altruistic, it does not require B to do so. However, a minor modification of the game ensures that B will become altruistic as well. Suppose that B has the option H not only of socializing at work but also of socializing at home.[16] I denote by z_2, i_2, c_2, and f_2 the quantities that correspond to z, i, c, and f when socializing occurs in both places. The payoffs are then

$$\begin{array}{cc} & \begin{array}{ccc} N & S & H \end{array} \\ \begin{array}{c} N \\ P \end{array} & \begin{bmatrix} 0, 0 & z, \quad -c & z_2, \quad -c_2 \\ 0, 0 & z - i, f - c & z_2 - i_2, f_2 - c_2 \end{bmatrix} \end{array}.$$

Suppose that the firm has lowered c sufficiently that $f - c$ and $f_2 - c_2$ are both positive, though the former is slightly bigger than the latter. Thus a selfish B would choose S if he expected A to choose P. I suppose also that i_2/f_2 is smaller than i/f. This means that the ratio of B's benefit to A's extra expenditure is particularly high when socializing takes place in both places.

Without altruism, both A and B choose N in equilibrium. But both players will choose to become altruistic when this option is open to them. Individual B will choose a γ_B slightly above $[(f - c) - (f_2 - c_2)]/[(z_2 - i_2) - (z - i)]$, and A will choose γ_A slightly above i_2/f_2. With these γ's, A chooses P and B chooses H. To see that this is an equilibrium, note that a lower γ_A leads B to expect A to play N, so B plays N as well and A loses. Similarly, a lower γ_B leads A to expect that B will, at best, play S so that A chooses N and B loses.

C. *The Evidence from the Hawthorne Experiments*

In the previous subsections we saw that incentives for group output as well as opportunities to socialize may facilitate the formation of altruism, which could then raise productivity. In this regard, the experiments at Western Electric's Hawthorne plant reported by Roethlisberger and Dickson (1939) are particularly relevant because they varied both the incentives under which groups of workers operated and their opportunity to socialize. My interpretation of these experiments is that output had the largest and most sustained rise when

[16] Nothing would change if he also had the additional option of socializing only in the home.

groups were given both strong incentives and opportunities to socialize. Either the opportunity to interact or strong incentives, when instituted one at a time in separate experiments, had a much smaller impact.

Before participating in the relay assembly room experiments, the workers received payments that depended on the aggregate output of a group of approximately 100 operators who were located in a large room. In the experiments, six workers, of whom five were assemblers and one was a layout operator, were separated into a small room and paid as a function of the output of this smaller group. In the experimental conditions, a worker who produced an additional unit thus received an additional one-fifth of the group's piece rate, whereas in the initial conditions, she received only one-hundredth. Not surprisingly, given that the payment per marginal unit rose, output rose as soon as the payment system was changed.[17] After this change in compensation, there were several experimental changes in the amount of rest the workers were allowed to take during the working day. During these rest pauses the workers socialized. Rest pauses were first introduced, and then their length and distribution over the day were varied. In one of the last experimental conditions there were no rest pauses at all. The principal finding of Roethlisberger and Dickson is that output increased permanently by about 30 percent after these changes (p. 160). This change in output is attributed by them to the fact that the five workers became friendly with each other and with their supervisors.

Roethlisberger and Dickson as well as many subsequent researchers have conducted further studies to elucidate whether the improved human relations in the workplace could really be responsible for the increased productivity.[18] Roethlisberger and Dickson conducted two additional experiments. In the second relay assembly experiment, the

[17] The figure on p. 56 shows what occurred in period 3.

[18] An interesting recent study by Jones (1990) argues that Roethlisberger and Dickson's detailed data on the output of each operator are consistent with this causal link. Jones shows that there is a predictable correlation between the outputs of different operators. Those operators who like each other are found to have individual outputs that are more highly correlated. If operators A and B like each other, then the lagged value of operator A's output is positively correlated with both operator A's and operator B's current output. Otherwise, it is positively correlated only with the current value of operator A's output. Jones thus shows that liking is related to interdependence in outputs. If the level of the operator's effort is indeed strategically complementary, then my model is consistent with this. With strategic complementarity, the marginal benefit to A from B's effort, α_b, is highest when A is working hard because α_{ab} is positive. Thus the increase in B's benefit from working harder when A raises his effort, i.e., the derivative of $\beta_j + \gamma_B \alpha_j$ with respect to a, is larger when B likes A and γ_B is high.

compensation scheme was changed and the experimental subjects were put in close proximity, but no other changes were instituted. This group's output also increased but only by 12.6 percent (pp. 129–32). This shows that the change in incentives was responsible for some but not all the increase in productivity.

In the final experiment, a group of workers originally paid on the basis of individual piecework were gathered in a room and given the same rest pauses and opportunities to interact as in the relay assembly room experiments. In this experiment, called the mica splicing room experiment, the compensation formula was the same in the original and the experimental conditions. Output first rose by about 15 percent but, contrary to what was observed in the relay assembly room, then declined so that by the end of the experiment it was only 4.4 percent higher than initially (p. 148).

Roethlisberger and Dickson emphasize the change in supervisory methods as explaining the change in productivity in the relay assembly room experiments. However, their own discussion of the mica test room experiment is more consistent with a model like the one presented in subsection *B*, which emphasizes the interpersonal relations between employees. In explaining why attendance was much more regular in the later stages of the relay assembly room experiment, they say that it was "more probable that the difference in attendance was due to a difference in attitude. . . . Among the Mica operators there was no pressure for attendance except that which came from the individual's own personal situation." In addition, "the Mica operators did not join in common social activities outside working hours. They had no 'parties' similar to the gatherings of the relay assembly girls. . . . There was no willingness to help one another. . . . The Relay Assembly Test Room was a 'group' story, the Mica Splitting Test Room was a story of 'individuals'" (pp. 155–56). The authors' own explanation is that "the relation between the type of payment system to the question of morale was overlooked. . . . Under group piecework, the operators in the Relay Assembly Test Room had a common interest under which they could organize. Under individual piece-work each girl was self sufficient; there was no need of working together" (p. 158).

What Roethlisberger and Dickson are suggesting is that not only the rest pauses but also the group piecework encouraged the friendship among the relay assembly room workers and that this friendship was in turn instrumental in raising output. To explain the increase in friendship and altruism, one needs to argue that the actions of the operators were strategically complementary. While most operators initially worked independently, there were at least two potential sources of complementarity. The first is that the layout operator had to help all assemblers. The marginal product of each assembler was thus a positive function of the layout operator's effort. Second, the assemblers found it possible to increase joint output by helping each other.

D. *Other Evidence on Cohesiveness and Productivity*

Stodgill (1972) surveys 34 studies of the connection between how much members of a group report liking each other (cohesiveness) and productivity. The results are quite mixed: about one-third of the studies find no relationship, one-third find a significantly positive relationship, and one-third find a negative one. A more recent study that also shows a trivial correlation of output with cohesiveness is Gladstein (1984). One problem with these studies is that they leave vague the participants' payoffs and, in particular, their compensation. I nonetheless explore here whether the mixed results are consistent with treating cohesive groups as altruistic. One experiment that is of particular interest because it obtains both negative and positive correlations between cohesiveness and productivity and whose results suggest that cohesiveness is indeed similar to altruism is reported in Berkowitz (1954).

Berkowitz studied the behavior of individual experimental subjects who believed themselves to be producing output together with a coworker (even though no coworker actually existed). Each subject's output was supposedly used by the coworker to make a finished product. The perceived production function thus exhibited strong complementarities between the effort of the two workers.

While first telling the subjects about their task, Berkowitz also induced favorable or unfavorable feelings for the coworkers. This was done by telling the subjects either that the coworkers were likely to be compatible with the subject or that no such compatibility existed. The subjects then briefly met their "coworkers" (who were just confederates of the experimenter). Those told that their coworkers were compatible did later express much greater liking for their coworker than those told that they were incompatible.

After the initial meeting, the subjects carried out their assigned task in isolation. During this period they received messages that they were told had been written by their coworkers. The key result of the study is that the subjects who expressed that they liked their coworkers were much more likely to raise their effort when these messages said that the coworker needed more inputs. They were also more likely to reduce their output if the messages said that the coworker either was tired or was inundated with materials. Thus workers who express that they like their coworkers in questionnaires also choose their output taking into account the desires of their coworkers. Thus cohesiveness and altruism are linked as in subsection *B*.

One question that arises at this point is whether the altruism pattern is consistent with my theory. Answering the question is made difficult by the lack of information on what the subjects believed to be their payoffs. If subjects felt that high levels of final output were beneficial, their effort was a strategic complement with the effort of their coworkers. Thus the theory predicts that the subjects should become altruistic toward their coworkers. The theory can also explain the lack of altruism between the "incompatible" coworkers if the subjects felt that their personal payoffs had a Prisoner's Dilemma structure as in Section I*B*. In that setup, there is no benefit from becoming altruistic toward someone who is not expected to reciprocate.

III. Relations of Authority

In this section, I study the relationship between employees in authority and their subordinates. I argue that, in contrast to what is predicted by the models of collusion in hierarchical agency (Tirole 1986; Holmström and Milgrom 1990), the current model predicts relatively little ganging up between superiors and their subordinates against the firm as a whole. The reason is that the strategic variables that they control do not tend to be strategic substitutes.

I consider two sorts of situations. In the first, the supervisor has some control over the wages of his subordinates. In the second, the supervisor has some authority but no ability to affect the employee's pay. I show that the supervisor is more likely to be altruistic in the latter case.

A. *Supervisors Who Determine Wages*

A particular example of a supervisor who can affect his subordinate's wages is provided by Tirole (1986). He considers the case of a supervisor with a single subordinate. At least sometimes, the supervisor has information that could be used to lower the subordinate's wage. If he wants to, the supervisor can withhold this information so that the subordinate's wage remains high.

The subordinate has to choose his effort. A higher level of effort tends to raise output and thus raises the supervisor's compensation. If one ignores any effect of this higher effort on the utility of the subordinate, the payoffs of the two employees have the following general form:

$$\begin{array}{c} \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cc} \quad C \quad & \quad D \quad \end{array} \\ \begin{bmatrix} f, z+g & 0, z+g \\ f, \ z & 0, \ z \end{bmatrix} \end{array},$$

where f, g, and z are positive. In this matrix, C represents the action that benefits the other player, whether it be withholding of negative information or high effort. In this game, each player's optimal action depends on his own altruism. But no matter what the level of his altruism, his optimal action is independent of the other player's expected action. Thus there is no strategic complementarity and no altruism in equilibrium.

In this model the supervisor neither gains nor loses from his altruism. When, instead of being able to affect the wage of only one employee, the supervisor can affect the distribution of wages among employees, altruism may actually be detrimental to the supervisor. When there are several subordinates, one important role for the supervisor is to maintain incentives for his subordinates by inducing them to compete with each other for the fixed amount of money he has available to compensate them. These incentives disappear if the supervisor becomes altruistic toward one employee.

Suppose that there are two subordinates, A and B.[19] I suppose that, as in Tirole (1986), the supervisor sometimes knows the true state of nature and sometimes does not. In particular, the supervisor sometimes knows the ranking of the two subordinates' outputs, whereas in other cases, he and the firm as a whole know only their sum.

With probability π, the firm and the supervisor observe only the sum so that they must compensate the two employees equally. With probability $1 - \pi$, the supervisor knows, and can prove to top management, which worker produced the most. To provide incentives for effort, it then makes sense to penalize the worker who has produced less and give a bonus to the worker who has produced more. This is generally better, for insurance purposes, than simply paying more to the more productive worker.[20]

[19] Laffont (1990) considers a related model with two employees who share one supervisor. He shows that if the supervisor can collude with one employee, then the firm will make less use of the supervisor's reports about his unverifiable information (because these reports are biased).

[20] This can be seen as follows. Suppose that we are considering a state of nature in which, in the absence of information, both workers would get w. Suppose that the contract then says that, with information on rank, the most productive worker gets $w + z$ and the other keeps getting w. With concave utility, one can always improve on this by offering a slightly higher wage when there is no information and a slightly lower wage when the employee comes in second. Suppose that this is done in such a way that, for fixed effort, expected payments are the same. So, if effort does not

I now ignore employee altruism to focus on the altruism of the supervisor for *one* employee. If the supervisor becomes altruistic toward one employee, he ceases to reveal his knowledge when his favorite is second. This eliminates the incentive effect of the contract on the employee who is not the favorite (since he never officially wins). The favorite still has some incentive to make an effort from the tournament component since he collects the prize only when his victory can be verified. Nonetheless, the reduction in the effort of the nonfavorite employee also reduces the favorite's effort. So output falls. If the supervisor's income depends positively on output, he is thus worse off by being altruistic toward one employee. Not surprisingly, favoritism can be quite detrimental for supervisors.

B. The Team Leader

In this subsection, I shall consider an authority relation based on Crozier's (1964) account of work in a French clerical agency. In this agency, the actual work was done by teams of four people whose job was to process customers' requests. The first member of the team checked the customers' records. The next two processed the requests that were approved by the first. The final member of the team checked the whole transaction. The first member was the team leader in the sense that she set the pace even though she did not have any formal authority or ability to set compensation (Crozier 1964, p. 18).

What makes this relationship interesting from our point of view is the difference in affective attitudes among the members of these groups. The employees had very few ties of affection with their coworkers. As Crozier says, "Only 20% feel positively about their workmates as potential friends" (p. 65). And most of these feelings of affection flowed from the team leaders to the other members of the team: "The middle-class girls who made positive comments about their workmates as possible friends were team leaders (half of them) and a few senior employees from the special workroom" (p. 36).

Given the connections between altruism and affect discussed in Section II, I analyze whether team leaders had an incentive to become altruistic. I thus present a model of a relationship between a team leader and a single other team member that is loosely based on Crozier's account. I assume that both employees can operate at one of two paces. In principle, the leader sets the pace. So she chooses first

change, the individual and the firm are indifferent to small perturbations of this sort. This perturbation raises the gap between the winner and loser so that the worker can make himself better off by raising his effort. This increase in effort benefits the firm as well.

whether to go fast (F) or slow (S). The other team member (whom I shall term the follower) then chooses whether to work fast (f) or slow (s). When both choose the slow pace, they both receive a payoff that I normalize to zero. When only one chooses a fast pace, the fast employee loses c and the other employee gets zero.[21]

The interesting case occurs when both choose the fast pace. I assume that the payoffs in this case depend on features of the environment that only the leader knows. In particular, there are circumstances in which the whole organization benefits from a fast pace. Teams that perform well during these circumstances might get promoted, might be favored when it comes to allocating better office equipment, and so forth. In this case the payoff from working fast is b for both employees (though only the leader currently knows this).

There are other circumstances in which only the leader benefits from a fast pace. These might be situations in which the leader's cost of handling the documents is particularly low. In this case, the leader gains d but the follower loses c. There might, in addition, be circumstances in which neither benefits from a fast pace, but this does not matter for the analysis. The employees are thus playing one of two games, which I assume to be equally likely. The payoffs in these games can then be written as

$$\begin{array}{c} \\ s \\ f \end{array}\begin{array}{c} \begin{array}{cc} S & F \end{array} \\ \begin{bmatrix} 0,\ 0 & 0,\ -c \\ -c,\ 0 & b,\ b \end{bmatrix} \end{array}, \quad \begin{array}{c} \\ s \\ f \end{array}\begin{array}{c} \begin{array}{cc} S & F \end{array} \\ \begin{bmatrix} 0,\ 0 & 0,\ -c \\ -c,\ 0 & -c,\ d \end{bmatrix} \end{array}.$$

When there is no altruism and $c > b$, the only equilibrium is $\{S, s\}$. The reason is that the leader would always play F if she felt that the follower would respond by playing f. But since the games are equally likely, the follower would then lose an average of $(c - b)/2$. The follower thus responds with s, which means that the leader is better off setting a slow pace.

It is apparent that the follower has nothing to gain by becoming altruistic. Such altruism will tend to induce her to respond to F with f, but this is not in her ex ante best interest. On the other hand, suppose that the leader becomes altruistic with a γ slightly greater than c/d and consider the game on the right. The leader now prefers a payoff of zero to having the follower choose f in response to F. The leader's altruism implies that she dislikes $\{F, f\}$ in the circumstances in which only the leader benefits from the fast pace.

Now suppose that such an altruistic leader expects the follower to

[21] In practice, matters were more complex. If the leader chose a fast pace, the follower's only recourse was to complain. These complaints were common and triggered investigations that were probably costly to both employees.

match the leader's actions so that she responds to S with s and responds to F with f. The leader would then choose F only when the game on the left is played (so that both benefit from the fast pace) and would choose S otherwise. Given these choices of the leader, it is indeed in the follower's best interest to match the leader's pace. The result is that both players gain an average of $b/2$ relative to the case in which both are selfish.

Even if she knows that the leader will be altruistic, the follower has nothing to gain by raising her γ. Altruism by the leader has a very particular rationale. Its benefit is that it leads the follower to believe that the leader will not move the team in a direction that benefits the leader at the expense of the follower. The result is that the follower trusts the leader. This trust then allows the leader to move the team in directions that are beneficial to all members.

IV. Sequential Moves

So far I have generally assumed that A and B select their actions a and b simultaneously. While somewhat artificial, this modeling device does ensure that A and B are treated symmetrically. In this section, I briefly consider the case in which A always chooses his action first. I do this both to study the robustness of my conclusions and to relate this work to that based on Becker (1974).

Consider the continuous action model of Section II, and assume that A chooses a first. Then B makes his choice of b with full knowledge of a. For any value of γ_B (including zero), the first-order condition that gives the resulting value of b is equation (6). Since a is known at this time, (6) gives b as a function of a and γ_B. Denote this function by $\Phi(a, \gamma_B)$. We know the local behavior of this function from (6), which implies that

$$db = \frac{\beta_{ab} + \gamma_b \alpha_{ab}}{\beta_{bb} + \gamma_b \alpha_{bb}} da - \frac{\alpha_b}{\beta_{bb} + \gamma_b \alpha_{bb}} d\gamma_b. \tag{17}$$

From (17) it follows that if a and b are strategic complements for both players, b rises when a does. Also, given the normalization that α_b is positive, b rises with γ_B.

Individual A's problem is different: he chooses a knowing that B will choose b using the function Φ. Thus a selfish A would set a such that

$$\alpha_a + \alpha_b \Phi_a = 0. \tag{18}$$

I now consider the initial stage in which the two agents choose their γ's to maximize their material payoffs. I assume that the two γ's are chosen simultaneously at this stage so that one agent's choice of γ

cannot influence the other agent's choice of γ. Thus A has nothing to gain from being altruistic. The reason is that changes in γ_A affect b only through the changes in B's perception of what A will do. But a is already known when B takes his action, so B reacts to a rather than to A's altruism. Therefore, A's material welfare is highest when a is chosen to maximize α itself. In other words, A achieves the highest value of α by keeping γ_A equal to zero.

This result—that any altruism that occurs in equilibrium must be B's—accords well with the timing of moves in Becker (1974, 1976), where the altruist moves last. It therefore suggests that Becker's order of moves need not be an assumption but a conclusion from letting altruism be endogenous. This result does hinge, however, on a predetermined order of moves.

I now study whether B benefits from becoming altruistic, that is, whether $d\beta/d\gamma_B$ is positive. From (17) and (6), $d\beta$ is given by

$$d\beta = \beta_a da + \beta_b db = \left(\beta_a + \gamma_B \alpha_b \frac{\beta_{ab} + \gamma_b \alpha_{ab}}{\beta_{bb} + \gamma_b \alpha_{bb}}\right) da + \frac{\gamma_B \alpha_b^2}{\beta_{bb} + \gamma_b \alpha_{bb}} d\gamma_B. \tag{19}$$

The second term represents the "direct" effect, which is negative given the second-order conditions. I shall consider the case in which γ_B is small so that this term can be neglected. The first term then has the same sign as da. To obtain da, one must differentiate (18) with respect to a, b, and γ_B. This raises the question of how Φ_a, which is the coefficient of da in (18), responds to changes in these variables. It is apparent from (18) that Φ_a depends not only on γ_B but also on b itself. Moreover, the extent of this dependence is a function of the third partial derivatives of α and β. Thus because the sign of these partial derivatives affects the desirability of altruism, strategic complementarity is clearly not enough. I thus proceed to show only that strategic complementarity plays *some* role in generating altruism. To do this, I assume that α and β are quadratic so that db/da varies only with γ_B. The derivative db/da with respect to γ_B, $\Phi_{a\gamma_B}$, is then

$$\Phi_{a\gamma_B} = -\frac{\alpha_{ab} + \alpha_{bb}\Phi_a}{\beta_{bb} + \gamma_b \alpha_{bb}}. \tag{20}$$

The second-order conditions imply that the denominator of (20) is negative. Thus $\Phi_{a\gamma_B}$ is positive as long as α_{ab} exceeds $-\alpha_{bb}\Phi_a$. In the normal case in which α_{bb} is negative, this requires a strong strategic complementarity.

If we ignore the dependence of Φ_a on b, differentiation of (18) gives the following expression for da:

$$da = -\frac{(\alpha_{ab} + \Phi_a\alpha_{bb})\,db + \Phi_{a\gamma_B}\,d\gamma_B}{\alpha_{aa} + \Phi_a\alpha_{ab} + \alpha_b\Phi_{aa}}. \tag{21}$$

When we combine (17) and (21), the change in a when γ_B is small is

$$da = \frac{-\alpha_b\Phi_{a\gamma_B}\beta_{bb} + \alpha_b(\alpha_{ab} + \Phi_a\alpha_{bb})}{D'}, \tag{22}$$

where

$$D' = (\alpha_{aa} + \alpha_{ab}\Phi_a)\beta_{bb} - (\alpha_{ab} + \alpha_{bb}\Phi_a)\beta_{ab}.$$

The denominator D' has a generally ambiguous sign. As before, this sign can be made determinate if one assumes that the system is stable. In particular, suppose that B always chooses b to maximize his utility (so that b satisfies [6]) but that A has to grope toward his optimal a.[22] If the change in a has the same sign as α_a, the system is stable when D' is positive.[23]

It is apparent from (20) that, with a positive D', the expression in (22) is positive whenever the strategic complementarities are sufficiently strong to make $\Phi_{a\gamma_B}$ positive. Then both terms in the numerator of (22) are positive so that a increases with γ_B. This in turn implies that increasing γ_B from zero is attractive since the increase in β then has the same sign as the change in a.

I have now established that strategic complementarities are also important in generating altruism in sequential move games. Before closing this section, I briefly discuss the application of these ideas to the "merit good" game of Becker (1991) and to the "transfer game" of Becker (1974). In the former, the first mover (the child) picks the level of a merit good such as studying hard or visiting often. After this, the parent, whose utility is increasing in the amount of the merit good picked by the child, chooses how much money to transfer to the child. Becker (1991) shows that children of altruistic parents will tend to pick relatively high values of the merit good if, from the point of view of the parent, transfers and the merit good are complements, that is, if the marginal utility of transfers goes up when the child behaves well. What my analysis shows is that precisely in this case in

[22] This makes some sense since A cannot be sure what b will be, whereas B does know a.

[23] To see this, note that (18) then implies that da/dt is proportional to

$$(\alpha_{aa} + \alpha_{ab}\Phi_a)(a - a^*) + (\alpha_{ab} + \alpha_{bb}\Phi_a)(b - b^*)$$

and use (17) to determine $b - b^*$ from $a - a^*$.

which altruism leads the child to behave well, altruism is in the parent's self-interest.

In the transfer game of Becker (1974), the action taken by the first mover A affects the income of both A and B. The second mover, B, then decides how much of his own income to transfer to A. To determine whether altruism is good for B in this game, we must determine whether the resulting increase in transfers to A leads A to generate more income for B in the initial stage.

I now argue that the existence of this strategic complementarity depends on whether A's action generates mostly income controlled by A or mostly income controlled by B. Whoever controls the income later decides how it is to be spent. Suppose first that A's action, a, requires costly effort to generate income only for himself and that b represents the fraction of B's income that is transferred to a. An increase in b then lowers a since it lowers A's marginal utility of income. The actions are strategic substitutes, and B loses from his altruism.[24]

Now suppose, as is implicit in Becker (1976), that A's action generates only income under the control of B. Then letting b be positive rather than zero raises a because now A gets at least a fraction of the income generated by his effort. Thus, in this case, a and b are strategic complements for A, and B does benefit from his altruism. His altruism leads A to produce, and he gets to keep a fraction of this income.

V. Conclusions

This paper has shown that altruism can be rational and that this can explain certain observations from industrial settings. The central idea is that strategic complementarity breeds altruism. By becoming altruistic toward you, I am led to change my behavior. If the resulting changes in actions induce you to change your own actions in a way that benefits me, my becoming altruistic is smart. This idea, while extremely simple, is rich in refutable implications. This richness comes from the fact that the theory predicts both the emotions that people have and their actions. Insofar as emotions can be measured, there are thus more implications than in theories that predict actions alone. Here, I have focused only on measurements of emotions that

[24] One interesting feature of this example is that the outcome is Pareto optimal (so that the rotten kid theorem holds) whether B is altruistic or not. The example is not wholly in the spirit of Becker (1974) because a affects only A's income. However, the conclusion that a and b are strategic substitutes for A so that altruism is bad for B extends also to the case in which the effort a also raises the income controlled by B as long as this effect is sufficiently small. Pareto efficiency of A's effort is still assured by the rotten kid theorem as long as A's utility is linear in income and B's altruism is sufficiently intense.

are based on self-reports and have found some concordance between these emotions and those predicted by the model. To be more convincing, this evidence would have to be supplemented with evidence from physical and chemical changes that are supposed to accompany changes in emotions.

In this concluding section, I deal with two issues. The first is a comparison of my theory of altruism with the idea of Homans (1950) and Thibaut and Kelley (1959) that people tend to like (or feel altruistic toward) those from whom they receive large rewards at low costs. The second is the applicability of the theory to other settings. I shall argue that many of the findings that have led Homans and others to embrace this idea are also consistent with my own. In addition, this traditional theory is less consistent than my own with the observed features of authority relations.

Using anthropological data from the Polynesian island of Tikopia and data from the Hawthorne experiments as evidence, Homans stresses that the friendliness among coworkers is much more pervasive than that between workers and their supervisors. On the island of Tikopia there is much more friendliness between brothers (who work together in the field) than between sons and fathers (the latter are the central authority figure). This seems problematic for the view that friendliness depends only on the frequency of interaction or even on the happiness of the outcomes. Indeed, Homans is led to posit that friendliness from repeated interaction arises only when the participants are relatively equal or, in his words, "when no one originates the interaction with much greater frequency than the others" (p. 243). One advantage of my theory is that this difference in affective attitudes is explained endogenously as one can see by comparing Sections II and III.

Moreover, my theory can explain *why* people feel altruistic toward individuals who give them large rewards (such as people who are competent or attractive). As I explain in Section II*C*, the benefit from this altruism is that it can credibly promise the recipient that he will get more benefits from their socializing. By the same token, rewards tend to be high whenever one is interacting with someone who feels altruistic toward oneself. The theory also explains why this altruism will often be reciprocated. The reason is that cooperative outcomes for *either* individual in the Prisoner's Dilemma obtain only when *both* individuals feel altruistic toward each other. This explains why feelings of altruism are often symmetric, as is required by the fairness theory of Rabin (1993).

In this paper, I have studied interactions only within the workplace. While unselfish regard for others certainly also takes place elsewhere, I have focused on the workplace because firms create a large degree of variation in the setting within which people interact. This allows one to study how relationships differ in different settings. The analysis of endogenous altruism might also allow one to understand other aspects of organizations. One remarkable feature of firms that has been stressed by Lazear (1989) is how rarely rank-order tournaments are used for purposes of compensation in spite of the fact that Lazear and Rosen (1981) have shown them to have desirable incentive effects. One reason for the paucity of such schemes may be that they encourage altruism insofar as a reduction in effort by one employee lowers the payoff to effort of his "competitors." Insofar as altruism rises in the presence of rank-order tournaments, they lose some of their desirable incentive effects.

While altruism in the work environment still deserves further study, extensions of the model to relationships among family members and friends also seem desirable. The endogenous creation of altruism within the family would seem to be a useful complement both to Becker's (1974, 1976) discussion of the consequences of altruism and to Becker's (1991) insight that people who like each other are more likely to marry.

Endogenous altruism from parents to offspring is considered by Mulligan (1993). He assumes that the level of altruism depends on the amount of time spent with one's children and shows that, as a result, parents with high wages end up less altruistic than parents with lower wages. It appears that one can reach a similar conclusion using my model even if one treats altruism as a choice variable that can be chosen independently of the time one spends with one's children. A different reason for poorer parents to spend more time with their children is that they do not have the resources to hire day care or to purchase other time-consuming forms of leisure. Given that they are going to spend more time with their children, it may be in parents' interest to become more altruistic because this leads their children to trust them more (as in the model of Sec. III*B*).

The paper has dealt exclusively with the altruistic ties among a small set of strategically related individuals. In practice, individuals are strategically linked to individuals, who, in turn, are linked to other individuals. There thus exist networks of strategic linkages that, according to my model, lead to networks of friendship and altruism. It would thus be interesting to know what the consequences of such networks are. Perhaps such networks can rationalize the view of Durkheim (1933) and Comte (1969) that the division of labor is the basis for "social solidarity."[25]

[25] Durkheim (1933, p. 373) says that "the economists would not have left this essential character of the division of labor in the shade . . . if they had seen that it is above all a source of solidarity"; Comte (1969, 4:478) says that "C'est donc le répartition continue des différents travaux humains, qui constitue principalment la solidarité sociale."

Appendix

Proof That $l > g$ Implies Cooperation in the Prisoner's Dilemma

Consider B's choice of action after the altruism parameters are set (so that the payoffs are given by [13]). If B believes that A will play C, his payoff is higher when playing C if and only if

$$1 + \gamma_B > 1 + g - \gamma l \quad \text{or} \quad \gamma_B > \frac{g}{1 + l}. \tag{A1}$$

If A plays C, B will play C if his γ_B satisfies this condition, and D otherwise.

If B believes that A will play D, his payoff is higher when he plays C if and only if

$$-l + \gamma_A(1 + g) > 0 \quad \text{or} \quad \gamma_B > \frac{l}{1 + g}.$$

If A plays D, B will thus play C if γ_B satisfies this condition, and D otherwise.

The issue is now whether $g/(1 + l)$, the cutoff γ_B for (A1), is larger than $l/(1 + g)$ or not. The former is larger if g exceeds l and is smaller otherwise. Thus if l exceeds g, there exist levels of γ_B between $g/(1 + l)$ and $l/(1 + g)$ such that B plays C if he expects A to play C; B otherwise plays D. Instead, if g exceeds l, any level of γ_B that incites him to play C when A plays C also leads B to play C when A plays D.

I shall now describe the Nash equilibria of this game. Consider first the case in which $l > g$. Suppose that both A and B have chosen γ's between $g/(1 + l)$ and $l/(1 + g)$. Then both will play C if they expect their opponent to play C, and they will play D if they expect their opponent to do likewise. Both $\{C, C\}$ and $\{D, D\}$ are Nash equilibria. However, both equilibria are not equally plausible since $\{C, C\}$ is Pareto dominant. So $\{D, D\}$ is not robust in the sense of Farrell and Saloner (1985). If A and B make their choices sequentially rather than simultaneously, the only subgame perfect equilibrium is $\{C, C\}$. I thus assume that $\{C, C\}$ is the only equilibrium in this case.

Now suppose that A sets γ_A to zero and B keeps his γ in this range. Then A will play D for sure so that B will choose to play D as well. The only Nash equilibrium is $\{D, D\}$. Note for completeness that if either player has a γ in excess of $l/(1 + g)$, then that player plays C regardless of his opponent's actions. The other player will react with C if his own γ is above $g/(1 + l)$ and with D otherwise.

I now argue that both A and B set their γ's between $g/(1 + l)$ and $l/(1 + g)$. Consider B's problem. If he sets γ_B in this range and A keeps γ_A equal to zero, B neither gains nor loses since the second-stage equilibrium is $\{D, D\}$, so that β is zero as it would be in the absence of altruism. If, instead, A also sets γ_A between $g/(1 + l)$ and $l/(1 + g)$, the equilibrium is $\{C, C\}$, so that β and α equal one. So setting γ_B in this range is a (weakly) dominant strategy for both B and A, and the equilibrium involves cooperation. While there does exist a Nash equilibrium in which the agents both remain selfish, this is, once again, not robust to the perturbation of Farrell and Saloner (1985), where the individuals pick their γ's sequentially. Note also that setting γ_B above

$l/(1 + g)$ is not as attractive because A would then prefer to remain selfish and receive a material payoff of $1 + g$.

I now turn to the case in which g exceeds l. If both γ's are less than $l/(1 + g)$, each agent plays D. If one player has a higher γ and the other player's γ is zero, the altruistic player chooses C and the selfish player chooses D. If both players set their γ above $l/(1 + g)$, then the outcome depends on whether their γ's are also above $g/(1 + l)$. If they are, the equilibrium has both A and B playing C. If at least one of the players has a γ between $l/(1 + g)$ and $g/(1 + l)$, then one player will play C and the other will play D. The reason is that a player with a γ in this interior range plays C in reaction to D but plays D in reaction to C.

When $g > l$, there is no altruism in equilibrium. If A sets γ_A above $l/(1 + g)$ (the minimum level such that the altruism affects his actions), then β is higher if B chooses to remain selfish. With γ_B equal to zero, β equals $1 + g$ instead of being at most equal to one if B becomes altruistic. Similarly, if A remains selfish, β equals zero if B remains selfish and equals $-l$ if he sets γ_B above $l/(1 + g)$. It is thus a dominant strategy for B to remain selfish, and the same argument applies to A.

References

Akerlof, George A. "Loyalty Filters." *A.E.R.* 73 (March 1983): 54–63.

Becker, Gary S. "A Theory of Social Interactions." *J.P.E.* 82 (November/December 1974): 1063–93.

———. "Altruism, Egoism, and Genetic Fitness: Economics and Sociobiology." *J. Econ. Literature* 14 (September 1976): 817–26.

———. *A Treatise on the Family.* Enl. ed. Cambridge, Mass.: Harvard Univ. Press, 1991.

Becker, Gary S., and Murphy, Kevin M. "A Theory of Rational Addiction." *J.P.E.* 96 (August 1988): 675–700.

Bergstrom, Theodore C. "A Fresh Look at the Rotten Kid Theorem—and Other Household Mysteries." *J.P.E.* 97 (October 1989): 1138–59.

Berkowitz, Leonard. "Group Standards, Cohesiveness, and Productivity." *Human Relations* 7 (November 1954): 509–19.

Bernheim, B. Douglas, and Stark, Oded. "Altruism within the Family Reconsidered: Do Nice Guys Finish Last?" *A.E.R.* 78 (December 1988): 1034–45.

Bulow, Jeremy I.; Geanakoplos, John D.; and Klemperer, Paul D. "Multimarket Oligopoly: Strategic Substitutes and Complements." *J.P.E.* 93 (June 1985): 488–511.

Camerer, Colin. "Gifts as Economic Signals and Social Symbols." *American J. Sociology* 94 (suppl., 1988): S180–S214.

Coleman, James S. *Foundations of Social Theory.* Cambridge, Mass.: Harvard Univ. Press, 1990.

Comte, Auguste. *Cours de philosophie positive.* 4 vols. Paris: Anthropos, 1969.

Crozier, Michel. *The Bureaucratic Phenomenon.* Chicago: Univ. Chicago Press, 1964.

Durkheim, Emile. *The Division of Labor in Society.* Translated by George Simpson. Glencoe, Ill.: Free Press, 1933.

Farrell, Joseph, and Saloner, Garth. "Standardization, Compatibility, and Innovation." *Rand J. Econ.* 16 (Spring 1985): 70–83.

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

Fudenberg, Drew, and Tirole, Jean. "The Fat-Cat Effect, the Puppy-Dog Ploy, and the Lean and Hungry Look." *A.E.R. Papers and Proc.* 74 (May 1984): 361–66.

Gauthier, David. *Morals by Agreement.* Oxford: Clarendon, 1986.

Gladstein, Deborah L. "Groups in Context: A Model of Task Group Effectiveness." *Administrative Sci. Q.* 29 (December 1984): 499–517.

Goodman, Paul S.; Ravlin, Elizabeth; and Schminke, Marshall. "Understanding Groups in Organizations." In *Research in Organizational Behavior,* vol. 9, edited by L. L. Cummings and Barry M. Staw. Greenwich, Conn.: JAI, 1987.

Hirshleifer, Jack. "Shakespeare vs. Becker on Altruism: The Importance of Having the Last Word." *J. Econ. Literature* 15 (June 1977): 500–502.

Holmström, Bengt. "Moral Hazard in Teams." *Bell J. Econ.* 13 (Autumn 1982): 324–40.

Holmström, Bengt, and Milgrom, Paul. "Regulating Trade among Agents." *J. Inst. and Theoretical Econ.* 146 (March 1990): 85–105.

Homans, George C. *The Human Group.* New York: Harcourt, 1950.

Jones, Stephen R. G. "Worker Interdependence and Output: The Hawthorne Studies Reevaluated." *American Sociological Rev.* 55 (April 1990): 176–90.

Laffont, Jean-Jacques. "Analysis of Hidden Gaming in a Three-Level Hierarchy." *J. Law, Econ., and Organization* 6 (Fall 1990): 301–24.

Lazear, Edward P. "Pay Equality and Industrial Politics." *J.P.E.* 97 (June 1989): 561–80.

Lazear, Edward P., and Rosen, Sherwin. "Rank-Order Tournaments as Optimum Labor Contracts." *J.P.E.* 89 (October 1981): 841–64.

Mayo, Elton. *The Human Problems of an Industrial Civilization.* New York: Macmillan, 1933.

Montgomery, James. "Revisiting *Tally's Corner:* Mainstream Norms, Cognitive Dissonance, and Underclass Behavior." Manuscript. Evanston, Ill.: Northwestern Univ., Dept. Econ., 1991.

Mulligan, Casey B. "On Social Determinants of Intergenerational Altruism and the Distribution of Wealth." Manuscript. Chicago: Univ. Chicago, Dept. Econ., 1993.

Rabin, Matthew. "Incorporating Fairness into Game Theory and Economics." *A.E.R.* 83 (December 1993): 1281–1302.

Raub, Werner. "A General Game-Theoretic Model of Preference Adaptations in Problematic Social Situations." *Rationality and Society* 2 (January 1990): 67–93.

Roethlisberger, Fritz J., and Dickson, William. *Management and the Worker.* Cambridge, Mass.: Harvard Univ. Press, 1939.

Smith, Adam. *Theory of Moral Sentiments.* London: Bohn, 1853.

Stodgill, Ralph M. "Group Productivity, Drive, and Cohesiveness." *Organizational Behavior and Human Performance* 8 (August 1972): 26–43.

Thibaut, John W., and Kelley, Harold H. *The Social Psychology of Groups.* New York: Wiley, 1959.

Tirole, Jean. "Hierarchies and Bureaucracies: On the Role of Collusion in Organizations." *J. Law, Econ., and Organization* 2 (Fall 1986): 181–214.