

Why Only Some Industries Unionize: Insights from Reciprocity Theory

By Sean Flynn^{*}

Dept of Economics
Vassar College
flynn@vassar.edu

Final Version: February 14, 2005
Forthcoming, Journal of Institutional Economics

Abstract

This paper argues that the degree to which a given industry's labor contracts are complete or incomplete is the major factor determining whether its workforce will be unionized. For instance, assembly line industries feature complete labor contracts because of the nature of the production technology: Either a worker keeps up with the line, or he does not. In such a situation, there is no chance for a reciprocal gift exchange under which firms offer high wages in exchange for high effort levels. The result is low wages that make workers prone to unionization. By contrast, jobs that feature incomplete contracts (lawyers, computer programmers, economists) already have reciprocity and gift exchange in place. Such benefits guarantee to workers that their better interests will be looked after by a management that wishes to maintain a positive and productive labor-management interaction.

1. Introduction

Why do only some industries unionize?

This question has never been given a good theoretical answer.

While labor-management relations were transformed radically during the 19th and early 20th centuries, no general theory exists to provide insight into the movement from craft guilds to large industrial unions that took place in the Industrialized Core during that

^{*} I would like to thank George Akerlof for kindly reading and commenting on an early draft of this paper, as well as Ernst Fehr and Matthew Rabin for discussing the idea with me. The comments and suggestions of two anonymous referees were also invaluable and greatly improved this paper. All errors are my own.

period. Neither has a general theory been developed that can explain why certain industries were relatively easy for unions to organize, while others remained union free. So far, only *ad hoc* hypotheses have been advanced. These discuss only how particular factors made particular industries at particular times prone to unionization. A general theory is needed.

I will argue that recent advances in our understanding of how reciprocity affects behavior can be used to explain which industries unionized, and I will show how that process was driven by the introduction of mass production, which eliminated the reciprocal labor-management contracting institutions that had existed previously. The key to this hypothesis is understanding how labor contract completeness affects the ability of workers and employers to engage in reciprocal gift exchanges of the type suggested by Akerlof (1982).

The importance of labor contract completeness was demonstrated in recent experiments by Fehr and Falk (1999). They show that reciprocity can express itself only if labor contracts are incomplete. Incomplete labor contracts allow workers to reciprocate high gift wages from employers with a gift of high effort levels. Fehr and Falk report that in experimental games with incomplete contracts, such behavior is the norm. Subjects acting as employers offer high gift wages, and subjects acting as workers reciprocate with high effort levels. Moreover, they do so even if the games are fully anonymous, meaning they have no reason to be nice simply to build up future cooperation.

By contrast, if labor contracts are complete, workers have no way of reciprocating a high gift wage with high effort levels. In otherwise identical games with complete contracts, wages offered by the subjects acting as employers are low. The employers apparently understand that complete contracting eliminates the ability of workers to reciprocate. As a result, the employers don't bother to offer high gift wages to elicit high effort levels.

These results are very interesting because if you look around the economy, you tend to find unions in industries where there are complete contracts. For instance, assembly line industries tend to be highly unionized. My hypothesis suggests that this is true because an assembly line worker has no chance to put forth higher effort or increase his output because the line moves at its own pace, independent of anything he may wish to do.

By contrast, unions are rare in industries where there are incomplete labor contracts that allow for a gift exchange relationship between individual workers and their employers. This is true for most professionals (lawyers, doctors, economists), as well as most service workers.

The hypothesis that labor contract completeness determines unionization also explains the decline of craft unions and the rise of industrial unions in the U.S. automobile industry. The key point is to realize that the change in union structure coincided with Henry Ford's switch to assembly line production methods.

Before the switch, cars had been assembled by skilled craftsmen working at their own pace in small shops. During that period, an incomplete labor contracting regime prevailed under which employers could offer high gift wages in exchange for the gift of high effort levels. The exchange of gifts kept workers happy and, as a result, union organizers had very little success organizing automobile workers.

The assembly line changed everything. It eliminated any chance for individual workers to reciprocate high gift wages with high effort levels. That, in turn, meant that firms no longer had an incentive to offer workers high gift wages in order to generate high effort levels.

This situation gave industrial trade unions the opportunity they needed. Workers who are not being kept happy by gift wages are susceptible to unionization. For Henry Ford and other industrialists, the shift to assembly line production was also a shift towards unionization.

The hypothesis I advance also suggests that you can view unionization as an institutional mechanism that can re-establish a gift-exchange relationship between workers and employers. While the complete contracting that comes with assembly line production precludes such a relationship between *individual* workers and employers, unions allow for the possibility of a gift-exchange relationship between the *collective mass* of workers and employers.

This point matters because it offers an alternative explanation for the fact that union wages are higher than non-union wages. The traditional view is that higher union wages are a type of rent, a payment in excess of productivity. I argue that high union wages may in some cases be the result of a gift exchange. Employers offer high gift wages that are reciprocated by unions that act to increase the productivity of the collective mass of workers.

Section 2 reviews Henry Ford's introduction of the \$5 daily wage at the same time he switched to mass production. It explains why Ford believed that such a large pay increase was necessary to prevent the unionization of his workforce in the wake of the switch to the assembly line. (Ford is used throughout the paper to demonstrate the applicability of reciprocity theory and labor contract completeness to the question of unionization.)

Section 3 reviews several labor market experiments that highlight the role of reciprocity between worker and employers. As these results are introduced, they will be related to various empirical observations that have been made about labor-management relations and about the behavior of workers under various labor contracting regimes. The key factor is shown to be the degree to which labor contracts are complete or incomplete.

Section 4 applies reciprocity theory to Henry Ford's decision to offer the \$5 daily wage. The case has already been convincingly made by Raff (1988) that Henry Ford offered that famously high wage to fight off unionization. The question that reciprocity theory

can answer is: Why did he have to fight off unionization at all? The answer is that the switch to assembly line production meant complete labor contracts that eliminated reciprocity between individual workers and Ford Motors. The section also argues that Ford later relented and stopped fighting unionization because he may have realized (consistent with experimental results) that it would be more profitable to pay lower wages and risk unionization.

Section 5 applies reciprocity theory more widely to other historical examples and notes that the introduction of more capital-intensive production methods in the late 19th century rang the death knell for craft unions but sounded the trumpet for the modern labor union. It also shows that reciprocity theory and labor contract completeness can be used to explain the low rate of unionization found among private sector service workers, as well as the high rate found among government sector service workers.

Section 6 concludes.

2. The \$5 Day at Ford Motor Company

In the winter of 1913-1914, Ford Motor Company switched to assembly line production. On January 5, 1914, Henry Ford raised the wages of his lowest paid employees from \$2.34 per day to \$5 per day. Skilled workers received proportionate increases.

Raff (1988) argues in great detail that Ford's decision to raise wages was a preemptive strike designed to prevent his workforce from being unionized. Raff argues convincingly that the sudden and large wage increase cannot be explained by other factors, such as changes in supply and demand in the labor market, a desire to reduce labor turnover costs, the need to reduce adverse selection problems, or the preemption of shirking.

Why did Ford Motor Company have a threat of unionization that it needed to preempt? I propose that the very nature of assembly line production made Ford Motor Company vulnerable to unionization. The vulnerability arose because the new production method destroyed the reciprocal labor-contracting institutions that had flourished before the assembly line. It is no coincidence that Henry Ford increased wages at the same time he switched to the assembly line. He realized that the new production method gave union organizers a much better chance of organizing his workers.

2A. Mass production implies complete contracts

To see how technology drove this process, it is best to look back at the automobile industry before Henry Ford introduced the "American System" of production. Ford Motor Company began in 1904 with a small shop of mechanics and artisans. The first product was called the A and was followed by the B, the C, and so on up to the famous Model T, which was the car that initiated mass production using assembly lines.

All Ford cars before the Model T had been handmade in small production runs. The workers involved in production during that period were skilled craftsmen. Industrial unions had tried futilely to organize them, but labor-management relations were handled through small craft unions that represented workers of the same occupation (core molders, for example). Furthermore, the labor contracts applied to such workers were by their very nature incomplete contracts. Wages were paid and output required, but no close supervision of employees was cost-effectively possible, as each worker was a craftsman working at his own pace to produce a quality product.

Production was vastly different on the assembly line. Each step in the production process required no great skill. A Ford supervisor bragged in 1914 that he could in a very short time “make a first class molder of a man who has never before seen a core-molding bench in his life.” Previous to this, core molding had been a skilled position, and the core molders had been represented by their own craft union.

In addition to eliminating the need for skilled workers, the assembly line also reduced the monitoring costs to zero. If a worker did not keep up with the pace of the line, inputs piled up at his workstation. Supervisors stationed on platforms overlooking the floor could immediately tell if a worker was not keeping up. When any laggard was spotted, he was, according to Martin (1915), immediately replaced with substitutes who “were constantly kept on hand, at factory’s expense, to meet all emergencies.”

The switch from shop production to assembly line production was thus a switch from incomplete labor contracts to complete labor contracts. In the former, input and output are only vaguely monitored. In the latter, input and output are perfectly monitored, and there is immediate verification of shirking or inability on the part of workers.

2B. Complete contracts eliminate gift wages and make unions more attractive

The switch from incomplete labor contracts to complete labor contracts also meant that it would no longer be possible to organize a gift exchange between individual workers and the employer. Because the assembly line moved at its own pace, an individual worker could not reciprocate a gift of a high wage with a gift of high output. That being the case, the employer was no longer motivated to offer a high gift wage to elicit high effort.¹

This point is crucial because gift wages are what appear to have staved off unionization under craft production. As long as workers were receiving high gift wages, they were happy and not interested in joining an industrial union. But with the shift to assembly line production, firms no longer had an incentive to give gift wages to motivate high effort. Thus, the new production technology itself appears to have been the entrée needed by industrial union organizers, who had tried unsuccessfully to unionize automobile

¹ But, as I argue below, a high gift wage could be used to elicit *other* gifts from workers. In particular, Henry Ford appears to have raised the wages of unskilled workers to \$5 per day as a gift designed to get them to resist union organizers.

workers throughout the craft production period. Assembly lines meant complete contracts, which in turn created a workforce disposed to union organization.

3. Reciprocity Theory and Labor Markets

The past 20 years have seen a profusion of experimental results that are inconsistent with the sort of selfish optimizing behavior that is presumed under standard, egocentric, utility-maximizing models of human behavior. The behavior perhaps most inconsistent with standard theory has to do with costly punishing. Players in experiments choose to punish players who they think are acting unfairly—even if applying the punishment is costly to the punisher, and even if games are structured so that players know there can be no future reward to punishing.

The most well-known example of this behavior comes from the ultimatum bargaining game, which was created by Güth *et al* (1982). In that game, the two players involved have to agree on the division of some amount of money X . The first player, A, makes a proposal to the second player, B, as to how to divide X . Player B has only two possible responses: He can accept the division proposed by A, or he can reject the proposal. If B rejects the proposal, then *neither* party receives any money at all.

Standard theory predicts that B will accept *any* proposal made by A because standard theory expects individuals to want to maximize their personal gain. Any positive amount is better than receiving nothing (which is what B would get if he were to reject), so standard theory predicts that all offers will be accepted.

Standard theory is wrong. A consistent result of the ultimatum game is that if the proposer A offers less than 30 percent of X to the responder B, then B rejects the offer. The standard theory fails because agents are interested in fairness and are willing to pay a cost—here, an opportunity cost—to avoid being treated unfairly.

Reciprocity theory attempts to capture the empirical regularity with which people reward positive treatment with positive responses and react to negative treatment with negative responses, even if those responses are costly and even if there is no long-term benefit to be had from revenge.² The behavior of agents in the ultimatum game affords an example of negative reciprocity. If A's proposal seems unfair to B, then B reciprocates negatively by rejecting the offer and seeing to it that A is harmed by getting nothing.

Note that rejection in the ultimatum game is not some sort of repeated-game punishment strategy. If B-type subjects think that they are being treated unfairly, they reject offers even if they know that they will never play the game again. Their rejection is motivated

² For definitions of reciprocity and evidence about reciprocal behavior in games, see Fehr and Gächter (2000), Cox and Deck (forthcoming), and especially Falk, Fehr, and Fischbacher (2000), which reviews and tests the existing fairness theories to show that both intentions and outcomes matter to people when they evaluate the fairness of a situation.

by a vindictiveness that is unrelated to any attempts to train the A-types to make fairer offers in the future.

People are positively reciprocal, too, and positive reciprocity can be seen very clearly in the trust game, first run by Berg *et al* (1995). In this game, a proposer receives X dollars from the experimenter. The proposer can then offer any amount Y from $[0, X]$ to the responder. The experimenter then takes Y and triples it, so that the responder actually receives $3Y$ dollars. The responder can then choose how much, if any, of that amount $3Y$ to send back to the proposer. Standard theory makes the bleak, sub-game perfect prediction that because a selfish, maximizing responder would never send any money back to the proposer, the proposer should never send any money to the responder.

Here, again, standard theory fails. When this experiment is conducted, proposers send money, and responders send money back. Moreover, there is a positive relationship between the amount of money sent by the proposer and the amount of money sent back by the responder: The more generous the proposer is to the responder, the more generously the responder reciprocates.

That is not to say, however, that money is sent over by all proposers or that money is sent back by all responders. Repeated experiments have shown that between 20 and 30 percent of subjects in trust games do not reciprocate. These subjects behave completely selfishly and according to standard theory. On the other hand, the same studies have found that between 40 and 66 percent of subjects behave reciprocally (see Fehr and Gächter, 2000). This paper focuses on how reciprocal behavior affects the propensity of workers to unionize.

The major point is that the introduction of mass production destroyed the previous labor-management relationship that had been based upon reciprocity and under which management had paid higher-than-minimal wages. Under that system, workers had taken those higher-than-minimal wages as a token of consideration, and they reciprocated by working with more-than-minimal effort. The assembly line ruined that system because it made both positive and negative reciprocity by individual workers impossible. There was neither any way to reciprocate positively by working better or faster, nor any way to reciprocate negatively by shirking.

But before applying reciprocity theory to the case of Ford Motor Company, let us review some experimental results on labor supply that greatly illuminate Henry Ford's decision to offer the \$5 daily wage. These results are important because they highlight the fact that reciprocal behavior can take place only under incomplete labor contracts. By contrast, with complete contracts, everything is specified, and there is no room for the contracting parties to behave reciprocally, either positively or negatively.

3A. The importance of incomplete contracts in generating gift wages

Akerlof (1982), Leibenstein (1987), and Akerlof and Yellen (1990) argue that labor markets are prone to gift exchanges in which employers offer pay wages in excess of

labor's opportunity cost, a behavior that elicits the reciprocal gift of higher-than-minimal effort on the part of workers.³ A key assumption underlying this line of thinking is that labor contracts are incomplete, so that poorly paid workers have the option of reducing their effort levels. In a complete labor contract, no such reductions are possible.

The critical effect that labor contract completeness or incompleteness has on wages is demonstrated by a series of experiments run by Fehr and Falk (1999). They show that wages are much higher under incomplete contracts—even in the presence of a large excess labor supply. The authors demonstrate this by running a competitive double oral auction to set wage rates, first in a treatment in which all contracts between workers and employers are incomplete, and then in a treatment in which all contracts between workers and employers are complete.

In both treatments, there are eight firms and 12 workers. Because each firm can employ only one worker, there is always an excess supply of labor, and you should consequently expect to see low wages. More specifically, the experiment was set up so that each worker faces a cost of contracting of 20. Consequently, with the given excess supply (12 workers but only eight jobs), the competitive wage level should be driven down to 20 under standard theory.⁴

Under the complete contract treatment, workers *had* to provide the contracted effort level because it would be monitored by firms that could impose stiff penalties for shirking. By contrast, the incomplete contract treatment was much like a trust game. After the wage contract was concluded, workers could choose *any* effort level they wanted because employers had no way to penalize workers who provided less effort than they promised when they contracted.

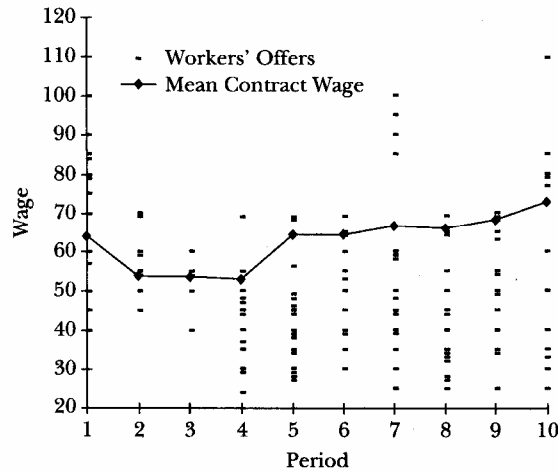
In either treatment, higher effort on the part of workers results in higher profits for the firm. This allows for the possibility of a mutually beneficial gift exchange in the incomplete contract case. Such a gift exchange would feature firms offering high wages in exchange for high effort. If such a gift exchange were to take place, you would expect to find higher wages in the incomplete contract treatment. However, given the huge excess supply of labor, standard theory would suggest that in either treatment competition between workers would drive wages down to their minimum level of 20. It is interesting to see once again how badly standard theory fails.

³ Efficiency wage theory can be divided into two streams of thought. The first was developed by Shapiro and Stiglitz (1984), Bowles (1985), Fehr (1986), MacLeod and Malcomson (1989) and others. Under this “shirking” version, there is a moral hazard problem, and firms pay high (incentive compatible) wages to prevent workers from shirking. The Akerlof (1982) version of efficiency wages may be termed the “fair-wage/effort” version. Under this version, if workers think wages are unfairly low, their effort falls. Consequently, raising wages can increase effort in some situations. Experimental evidence supports both versions of the efficiency wage hypothesis (see Fehr and Gächter 2005).

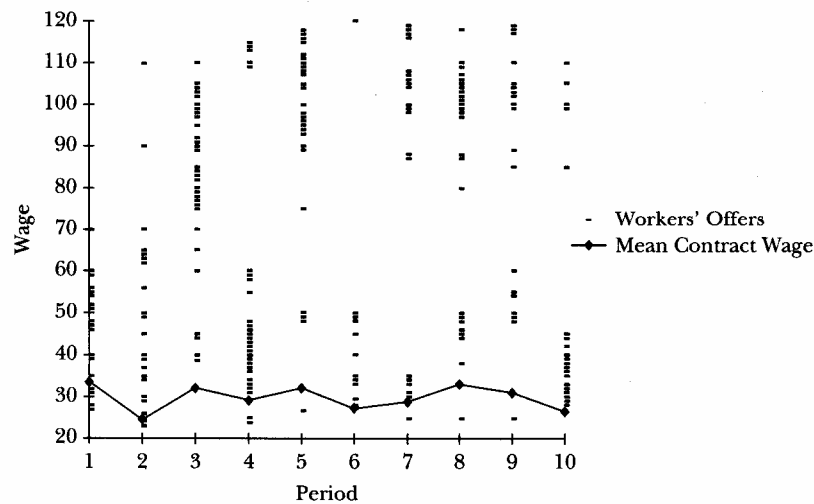
⁴ Individual workers would be indifferent to a contracted wage of 20 because their wage of 20 would just equal their cost of working of 20. Therefore, they would gain from any contracted wage higher than 20 and would lose if they contracted for a wage less than 20. Given these facts and the excess supply of labor, one would expect wages to be driven down to 20.

Figure 1

Workers' Offers and Mean Contract Wages in the Double Auction Market with Incomplete Contracts



Workers' Offers and Mean Contract Wages in the Double Auction Market with Complete Contracts



Source: Fehr and Falk (1999).

Figure 1 shows that average wage levels are much higher under the incomplete contract regime than under the complete contract regime. In addition, Figure 1 also shows the wage offers proposed by each of the workers in each period. As you can see, employers

tended to take the lowest offer under the complete contract regime, but chose to accept offers substantially higher than the lowest offer under the incomplete contract regime.⁵ The result is that incomplete contracts are associated with much higher equilibrium wage levels. This appears to happen because incomplete labor contracts allow positive reciprocity to manifest itself.

Under a complete labor contract, workers cannot take revenge against their miserly employers by giving poor effort. Under an incomplete contract, however, workers *can* take such revenge if they feel that they have been treated unfairly. The upshot is that firms are reluctant to accept a low wage bid. Firms operating under an incomplete contracting regime do not accept the lowest wage offers because doing so would mean hiring workers who would likely be disgruntled and not put forth high effort.

Instead, firms choose to pay high wages. A gift exchange takes place under which firms pay higher-than-minimal wages, and in return workers provide higher-than-minimal effort. Interestingly, the net effect is that firm profits are higher in the incomplete contracting case than in the complete contracting case. The higher level of output produced by happy, well-paid workers more than makes up for paying them high gift wages.

3B. Higher profits call for higher gift wages

By raising wages above the minimum level of 20 in Figure 1, incomplete contracts generate gift wages for workers. It turns out that the size of gift wages varies directly with firm profitability. This is demonstrated by Fehr, Gächter, and Kirchsteiger (1996), who present results similar to those above, except that the various firms offering employment have differing profit opportunities. The authors find a strong positive correlation between profit opportunities and the size of the gift wages paid to workers.

While this result is not compatible with competitive theories of wage determination, it can easily be explained as a result of gift exchange taking place under incomplete contracts. As argued by Akerlof (1982), whether a given wage is perceived as generous enough to be counted by workers as a gift depends on the circumstances. In particular, workers in very profitable firms feel entitled to share in the companies' success. Consequently, successful firms have to pay their workers higher gift wages than do less successful firms.

This hypothesis is useful in explaining the well-documented fact that workers in profitable firms and industries are paid more than workers in similar jobs in less successful firms and industries, even after taking account of differences among workers in terms of skills and human capital (see Dickens and Katz 1987 and Krueger and Summers 1988).

⁵ Moreover, people seem to know this implicitly. In Fehr, Klein, and Schmidt (2000), when subjects acting as firms were given the choice of whether to offer a complete or an incomplete contract, they almost always chose to offer the incomplete contract--despite the fact that they knew nothing of previous experimental results that showed that firm profits are higher under incomplete contracts.

The hypothesis is also relevant for analyzing Henry Ford’s decision to so precipitously raise wages in January, 1914. The reason is that Ford Motor Company was by far the most profitable firm in its industry, as you can see by looking at Table 1, which reproduces Table 3 of Raff (1988) and gives automobile manufacturers’ profits both annually and weekly circa 1913, about the time that Henry Ford decided to increase wages.

The size of Ford’s profits gives some hint as to why Ford felt that he had to so dramatically raise wages in order to keep his workforce resistant to unionization. He seems to have understood that his wage hike had to be quite large if it was to be perceived as a gift. Thus, reciprocity theory can explain not only why Ford felt the need to offer a gift wage to prevent unionization, but also why that gift wage was so large.

Table 1		
Profits of U.S. Auto Makers, circa January 1913		
Company	Annual Profit	Weekly Profit
Ford	\$27,087,204	\$541,744
General Motors	8,184,054	183,911
Willys-Overland	5,864,858	131,911
Packard	2,364,568	53,136
Studebaker	1,905,413	42,818

Source: Raff (1988)

3C. Monitoring for shirking ruins reciprocity

I argue above that reciprocity can explain experimental results that traditional self-interest theory cannot. Furthermore, when it is allowed to operate in an incomplete contract environment, reciprocity can lead to outcomes much better for *both* workers and firms than those that would prevail under complete contracts. In particular, reciprocity means that workers can get high gift wages, and firms can get high profits (thanks to the high effort levels provided by workers pleased by their high gift wages).

It might be imagined that explicit disincentives could also be used to motivate workers. However, Fehr and Gächter (2000) find that when disincentives are added to an incomplete contract game in which reciprocity is possible, the presence of disincentives by itself destroys reciprocal behavior. To highlight the effect of disincentives, they play two games that are identical except that one allows employers to punish workers for shirking.

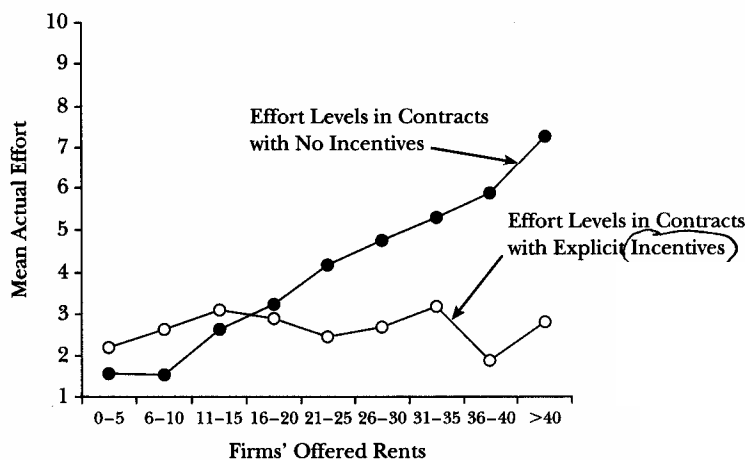
The game without punishment is an incomplete contract game in which firms offer a wage contract that requests that the worker expend some level e of effort. If a worker accepts the contract, however, there is nothing to prevent the worker from giving only the minimal effort $e_{low} = 1$ because workers can't be punished in this version.

As with other incomplete contract games, reciprocal behavior was observed when this game was played. Figure 2 (which reproduces Figure 3 of Fehr and Gächter 2000) gives the results of the game in the solid black dots. The figure shows that the higher the wage offered by firms on the horizontal axis, the higher the effort level with which employees reciprocated on the vertical axis.⁶

⁶ The horizontal axis is labeled "Firms' Offered Rents" rather than "Firms' Offered Wages" because Fehr and Gächter (2000) measure the amount by which wages exceed the minimum that would be necessary to get workers to give only the minimum effort level e_{low} . They refer to any wage in excess of this minimum as a *rent*. I reserve the word *rent* for situations in which workers are being paid more than their marginal product. Because gift wages can sometimes elicit an increase in productivity that is large enough to more than pay for the gift, gift wages are not necessarily rent.

Figure 2

Actual Effort-Rent Relation in the Absence and Presence of Explicit Performance Incentives



Source: Fehr and Gächter (forthcoming).

Behavior is very different in the version of the game that allows punishment. The punishment mechanism used in the game allows employers to stipulate in the wage contract a fine that will be paid if workers are verified to have shirked (by providing less than the effort level e desired by the firm). In addition, all parties know that the probability of verification given shirking is one-third.

As can be seen by the hollow circles in Figure 2, allowing employers to punish shirkers destroys reciprocal behavior. Workers no longer react to higher wage offers by putting forth higher effort levels.

To explain this result, Fehr and Gächter (2000) speculate that explicit disincentives “may cause a hostile atmosphere of threat and distrust, which reduces reciprocity-based extra effort.” What is even more striking is that reciprocity is destroyed by just a one-third chance of getting caught when shirking. This probability is much less than the nearly 100 percent chance of getting caught when shirking on a Ford assembly line. Indeed, it offers a second reason why you should not expect reciprocity when there are complete labor contracts.

The first reason was given above: The assembly line made reciprocal behavior impossible for individual workers because there was no way for any one worker to improve his output quantity or quality to reciprocate for a gift of high wages. To that result must now be added the conclusion that even had the individual worker been able to reciprocate in some way, he likely would have chosen not to do so because the assembly line provided Ford Motor Company with highly effective ways to monitor and thereby penalize underperformance.

4. Reciprocity and the \$5 Day at Ford

Raff (1988) demonstrates that preempting unionization is the only logical motive that can explain why Ford raised the pay of his workforce so drastically. Having reviewed the relevant experimental results about reciprocity theory, incomplete contracts, and contract disincentives, we are now in a position to elaborate on why Henry Ford felt the need to offer a \$5 daily wage in order to preempt the unionization of his workforce. These ideas can also explain why Ford Motor Company was the only firm in its industry that needed to undertake such a policy, why such a policy proved successful at preventing unionization, and even why the policy was eventually abandoned.

4A. The assembly line is a complete contract

Why was Ford Motors uniquely vulnerable to unionization and in need of taking preemptive action against unionization? Reciprocity theory suggests that Ford was uniquely vulnerable in 1914 because it was the only car maker to have switched to assembly line production. It was thus the only firm in its industry to have destroyed the previous incomplete contract labor regime. In its place, assembly line production created complete contracts that precluded the possibility of reciprocal behavior between Ford Motor Company and individual Ford employees.

The end of reciprocity resulted from two changes. First, a line worker would expend only the minimal effort needed to keep up with the line. If he worked any less hard, he would be fired. Attempting to work any harder would have been futile, as he could affect neither the quantity nor the quality of the final product by any personal effort. Consequently, Ford Motor Company had no incentive to pay gift wages in order to get any particular employee to expend more effort on his job. Because workers could not reciprocate, there was no point in giving a gift wage to encourage higher effort.

Second, the technology of the assembly line brought with it low monitoring costs and a nearly 100 percent chance of catching shirking workers. As demonstrated by Fehr and Gächter (2000), monitoring workers ruins reciprocity. But Ford had no choice. Monitoring *had* to be done. Ford supervisors hawkishly watched every station on the line because if even one station made an error or moved too slowly, the whole line could come to a halt. The result was a high-monitoring labor environment inconsistent with reciprocity.

4B. Preempting unionization by paying a *group* gift wage

With the death of reciprocity between the employer and individual workers, Ford Motor Company became uniquely vulnerable to unionization. Ford's competitors, all of which were still using craft workshops to produce automobiles, could rely on reciprocity to keep their workers happy. They could pay high wages to individual workers and expect that those individual workers would reciprocate by expending a great deal of effort in producing a good product. Firms could also expect that union organizers would be

ignored by happy, positively reciprocating workers. *If unions exist to raise wages, then they are of little benefit to workers who are already receiving high wages thanks to reciprocal gift exchange.*

This explains why raising wages to \$5 per day was a good way to prevent unionization. Doing so established a new gift-exchange relationship to replace the one that had been destroyed by the adoption of the assembly line. In exchange for the generous gift of high wages, workers would reciprocate by ignoring union organizers. After all, union organizers could hardly argue that they could raise wages much higher than Henry Ford had already done.

The key point, however, is that reciprocal relationships between an individual worker and the firm were not possible under the assembly line method of production. *If any gift exchange were to take place, it had to take place between the firm and the workers as a group.* That is why wages were raised across the board, for all workers.

Finally, it must be reiterated that Ford Motors had a unique problem. This can be gathered from the fact that after Ford Motors raised wages, no other car manufacturer followed suit. That is, not one of the competitors still utilizing craft-based production methods felt any need to raise wages.

4C. The eventual unionization of Ford

Henry Ford did not maintain a policy of high gift wages in the years after 1914. In fact, wages were barely raised at all in the coming decade, such that by the mid 1920s, union organizers were making steady progress at Ford's factories. Given that Ford was originally willing to pay high gift wages to prevent the unionization of his workforce, why did he not continue with this policy?

Quite simply, Henry Ford probably figured out another lesson of the experiment run by Fehr and Gächter (2000) that isn't visible in Figure 2: Although worker effort is lower in the complete contract game, profits are higher because the cost savings from not having to pay high, gift-exchange wages more than makes up for the reduction in output due to low effort. Stated a bit differently, while it was the case that workers in the incomplete contract treatment did give higher effort for higher wages, those increases in effort weren't nearly enough to justify the higher wages used to bring them about.

Ford may have discovered over time that the potential costs of dealing with a disgruntled workforce (to whom he did not pay gift wages) were less expensive than the costs involved in paying the gift wages necessary to keep unions out of his factories. This is especially likely to have been the case given the low-cost monitoring available under assembly line production and the fact that because the speed of the line was fixed, workers could not respond to Ford's 1914 gift wage by increasing production. As a

result, Ford's gift wages necessarily reduced Ford's profits. Ford most likely ended the policy of gift wages when he figured this out.⁷

5. Explaining Which Industries Unionize

Above, I discuss how reciprocity and labor contract completeness can be used to explain Henry Ford's decisions to implement and then abandon a policy of high gift wages. I will now argue that reciprocity and labor contract completeness can give a general explanation of why only certain industries unionize.

The key appears to be that only those industries in which it is difficult to establish a gift-exchange relationship between employers and *individual* workers are prone to unionization. Three broad categories of jobs seem to fit this description:

- Jobs in industries like automobile manufacturing where reciprocal gift exchange is precluded because of the nature of the production process.
- Government jobs in which the gift-exchange relationship is not available because all workers in a group are paid essentially the same and promoted essentially the same without regard for ability or exertion.
- Jobs in acting and construction, two industries in which output is instantly visible and is therefore subject to complete monitoring.

I will discuss each case in its own subsection.

5A. Some production processes preclude reciprocity

Figure 1 shows that the high wages that accrue to workers under an incomplete contracting regime disappear when a switch is made to complete contracts. Because unions are a mechanism by which high wages can be restored, you should expect to find them in situations in which a production process precludes reciprocal gift exchange and the high wages that it generates.

Aside from assembly line industries, piece rate industries are perhaps the best examples of complete contracts leading to high rates of unionization. For example, you see high rates of unionization in the garment industry where seamstresses are paid *per shirt made*, as well as in the mining industry, where workers are paid *per ton mined*.

The reason union organizers do so well in these industries is because piecework precludes reciprocal gift exchange between individual workers and employers. Once the piece rate is set, workers can only produce more or less. They cannot offer higher quality

⁷ Ford likely chose to end the policy gradually by not raising wages for a decade (rather than immediately cutting wages down to non-gift levels) because to have cut wages drastically down to non-gift levels would have immediately antagonized his workforce and made it prone to unionization.

to the employer. And there is no way for workers to modulate effort in a way that will hurt employers without hurting themselves, because any reduction in output only lowers the worker's own wages. Piece rates are therefore complete contracts that destroy the prospects for reciprocal gift exchange. And without gift exchange keeping workers happy, they are prone to unionization.

By unionizing, workers in such industries may be able to obtain higher wages. But if higher wages are achieved, should they be considered gift exchange or rent? The answer depends upon whether the higher wages are met with higher productivity.

Recent work suggests that, in at least some cases, unions can indeed raise productivity in exchange for higher wages. Notably, Black and Lynch (2001) find, in a representative sample of U.S. firms that unionized, firms that adopted joint decision-making coupled with incentive-based compensation had higher productivity than non-union plants that undertook similar policies. Similarly, Hamilton et al (2003) show that a unionized firm that switched from individual piece rates to group piece rates showed a 14 percent increase in productivity.

On the other hand, a recent meta survey by Doucouliagos and Larouche (2003) of papers investigating the relationship between unionization and efficiency found no overall effect when examining 79 studies covering eight countries. Consequently, the hypothesis advanced in this paper can confidently explain only how the absence of a gift exchange between *individual* workers and employers can lead to unionization. Whether the union can facilitate an aggregate gift exchange of higher pay for higher output between the workers *as a group* and the firm is less certain. However, given the large difference in productivity between incomplete and complete labor contracts, such a possibility is tantalizing.

5B. Why private service sector workers aren't very unionized

In the latter half of the 20th century, the fraction of the U.S. labor force that was unionized steadily declined.⁸ This decline has been ascribed to the ongoing movement of GNP away from manufacturing industries towards the service sector: As workers moved from the heavily unionized manufacturing sector towards the lightly unionized service sector, the overall fraction of unionized workers necessarily declined.⁹

But this explanation begs the question: Why are service industries less unionized than manufacturing industries?

⁸ Only 13.5 percent of all U.S. workers were union members in 2000. This represents a very large decrease from a high of almost 36 percent in the early 1950s.

⁹ In 1950, 45 percent of the non-agricultural labor force was employed in the service sector, while 65 percent was employed in the goods-producing sector. In 2001, the service sector accounted for more than 80 percent of the non-agricultural labor force, while the goods-producing sector employed less than 20 percent.

I propose that the difference in unionization rates between service industries and manufacturing industries can be ascribed to the fact that most labor contracts in service industries are incomplete labor contracts, while most labor contracts in manufacturing industries are complete labor contracts. The incomplete labor contracts found in service industries lead to reciprocal gift exchanges that keep wages high and workers happy. By contrast, the complete labor contracts found in manufacturing industries preclude the high wages generated by reciprocal gift exchange. The result is that workers in manufacturing industries are much more willing to join unions than workers in service industries.

5C. Why government service sector workers are heavily unionized

The major exception to the lack of unionization in the service sector is the high rate of unionization among government workers in the United States. Because nearly all U.S. government employees are service workers, the 37.5 percent unionization rate found among government workers would at first glance appear to be totally inconsistent with the low unionization rates found among service workers in the private sector.

It would also appear to contradict my contention that unions tend to arise only when there are complete labor contracts. After all, don't government workers have incomplete labor contracts like most other service sector workers?

I would say no. The civil service rules that cover government jobs create what is essentially a complete labor contract. Indeed, working for the government is a lot like doing piecework because promotion and pay increases are largely independent of effort.

For instance, under the compensation schemes used in public schools, the best high school teacher is paid the same as the worst if both have worked for the same number of years. Similarly, the firefighter in the neighborhood with the petroleum refinery is paid the same as the firefighter across town in the neighborhood where the buildings never burn.

Such effort-independent wage payments preclude gift exchange between the government and its individual employees. Without the ability to tailor pay to individual output, government managers cannot offer individual workers the gift of a higher wage in exchange for greater effort.

Consequently, I think that it's not surprising that service unions have had much better luck organizing public sector service workers than they've had organizing private sector service workers. While managers of service firms in the private sector can engage in gift-exchange incomplete contracts with individual workers, government managers, required by law to pay the same wage to huge numbers of workers despite manifest differences in ability and competence, cannot engage in gift exchange. The result is that government service workers are much more prone to unionization.¹⁰

¹⁰ As pointed out by an anonymous referee, this explanation for the high rate of unionization in the case of government service workers must be placed in context by noting that governments are often much less averse to seeing their workforces unionize than are private sector employers. The theory presented here is

It is interesting to note that under my hypothesis, the civil service protections given to government workers are the cause of the unionization of government workers. Along these lines, one can understand the great protests made by unions whenever plans are announced to outsource government jobs to the private sector.

Any time such a plan is announced, union organizers complain bitterly. This paper suggests that they are so vociferous because they realize that if these jobs are moved to private-sector firms, which are free to engage in incomplete labor contracts, workers doing those jobs will be much less likely to join a union. By contrast, if the jobs stay in the governmental sector, they will continue to have the civil service protections that make it extremely difficult to maintain a gift exchange between individual workers and government managers. It is for this reason that unions complain so loudly when service jobs are transferred from the government to the private sector.¹¹

5D. All the world's a stage: Full monitoring for construction workers and actors

Construction and performance are the two private-sector industries where unionization rates remain high in the United States. Construction workers, be they painters, carpenters, bricklayers, or steelworkers, are organized into guilds and unions that are very strong and able to demand very high wages. Similarly, actors and musicians have been organized into effective unions since ancient times,¹² and today every sort of actor and musician in the United States is compelled to join the appropriate union or be without work.

What both of these occupations share is immediate and full monitoring. This is most obvious in the case of performers. If they are not good, they will be immediately recognized as such, and they will not get paid. Quite simply, they do not have the chance for a reciprocal contracting relationship because of the complete monitoring that goes along with their jobs. Whereas workers in most jobs can choose to shirk and expend low effort, performers cannot.

The situation is similar for construction workers. Any bricklayer who does not lay straight rows of bricks will be fired immediately, as will any carpenter who cannot make a joint square. The labor contract is complete, and unions are a natural response.

6. Conclusions

a contribution because it gives an intuition for why government workforces may be more prone to unionization than private sector workforces, independent of their respective employers' feelings about unionization.

¹¹ A similar thing happened in the private sector at Harvard a few years ago when the university decided to outsource food service work and janitorial services. A great union-led outcry arose because university employees are easier to organize than private sector workers.

¹² Pliny the Elder reports that musicians and actors in ancient Rome and Greece were organized into guilds demanding minimal daily payment rates.

This paper argues that workers laboring under a complete labor contract will be more receptive to unionization than are workers laboring under an incomplete labor contract. The difference stems from the fact that gift exchange between workers and employers is possible only under incomplete labor contracts.

Under incomplete labor contracts, employers offer high gift wages that are reciprocated by workers offering high effort levels. Because reciprocal gift exchange keeps wages high and workers happy, unions have little to offer workers in such situations, and unionization rates are consequently very low.

By contrast, gift exchanges are not possible under complete labor contracts. Wages are much lower, and workers are much less happy. The result is that unions have much greater success organizing workers in industries that feature complete labor contracts.

The history of unionization at Ford Motor Company supports the hypothesis that labor contract completeness is the major factor determining whether or not workers will want to unionize. Before the assembly line, production was done by skilled craft workers working under incomplete labor contracts. This allowed for a reciprocal gift exchange that kept wages high and workers happy. As a result, industrial unionization was non-existent despite the best attempts of union organizers.

As shown by Raff (1988), however, when Henry Ford decided to switch to assembly line production, he also felt the need to offer an extremely high \$5 per day wage to prevent unionization. This paper argues that the reason he offered the wage increase at the same time that he switched to assembly line production was because he realized that the new production method involved a complete labor contract. Assembly line workers could not reciprocate a gift of higher wages with higher effort because the speed of the line was fixed.

Ford apparently understood that the new production method would destroy the previous incomplete contracting regime and make his workforce prone to unionization. The \$5 daily wage was Ford's preventative strike. If he paid high wages on his own, union organizers couldn't promise much improvement.

The decline of unions in the second half of the 20th century can also be explained in terms of complete and incomplete labor contracts. As workers moved from manufacturing jobs that featured complete labor contracts and high rates of unionization to service sector jobs that featured incomplete labor contracts and low rates of unionization, the overall rate of unionization in the economy fell.

In addition, the high rate of unionization found among governmental service sector employees can be explained as the result of civil service protections preempting the opportunity for reciprocal gift exchange. Because promotion and pay increases are tied to tenure rather than performance, government managers cannot exchange gifts of high wages for gifts of high effort on the part of workers.

Finally, the high rate of unionization among construction workers and performance artists can also be explained as a result of complete contracts. Reciprocal gift exchange is stifled in both industries by the immediate and complete monitoring of output.

If complete contracts are indeed the major factor leading to unionization, then unionization may offer a tantalizing way to obtain the productivity gains that accrue under reciprocal gift exchange. In situations like assembly line production where the production technology itself involves a complete contract between individual workers and the employer, there is no chance for productivity-enhancing reciprocal gift exchange. If a union is formed, however, the situation is completely transformed because the relationship between the union as a whole and the employer is that of an incomplete contract, in the sense that the union can shirk by striking or offer to raise productivity in exchange for concessions. In such a situation, a renewal of reciprocal gift exchange between the union as a whole and the employer may be possible.

This is no small matter given that productivity in experimental games can be so much higher under incomplete contracts than under complete contracts. Indeed, the switch from complete contracts (between individual workers and employers) to incomplete contracts (between unions and employers) that takes place when firms unionize may help to explain why unionization appears to raise firm productivity in some real world cases. Further research along these lines is needed.

Bibliography

- Akerlof, George, 1982, "Labor Contracts as Partial Gift Exchange," *Quarterly Journal of Economics* 97, pp. 543-569.
- Akerlof, George, and Yellen, Janet, "The Fair Wage-Effort Hypothesis and Unemployment," *Quarterly Journal of Economics* 97(4), pp. 543-569.
- Berg, Joyce, Dickhaut, John, McCabe, Kevin, "Trust, Reciprocity, and Social History," *Games and Economic Behavior* 10, pp. 122-142.
- Black, Sandra E., and Lynch, Lisa M., 2001, "How to Compete: The Impact of Workplace Practices and Information Technology on Productivity," *The Review of Economics and Statistics* 83(3), pp. 434-445.
- Bowles, Samuel, 1985, "The Production Process in a Competitive Economy: Walrasian, Neo-Hobbesian, and Marxian Models," *American Economic Review* 75, pp. 16-36.
- Clark, Gregory, 1984, "Authority and Efficiency: The Labor Market and the Managerial Revolution of the Late Nineteenth Century," *Journal of Economic History* 44 (December), pp 1069-1083.
- Cox, James, and Deck, Cary, Forthcoming, "On the Nature of Reciprocal Motives," *Economic Inquiry*.
- Dickens, William, and Katz, Lawrence, 1987, "Inter-industry Wage Differences and Firm Characteristics," in *Unemployment and the Structure of Labor Markets*, Lang, Kevin and Leonard, Jonathan, eds., Basil Blackwell Publishers, Oxford, UK.

- Doucouliafos, Chris (Hristos), and Larouche, Patrice, 2003, "What do Unions do to Productivity?" *Industrial Relations* 42, pp 650-691.
- Falk, Armin, Fehr, Ernst, and Fischbacher, Urs, 2000, "Testing Theories of Fairness-- Intentions Matter," *University of Zurich Working Paper* No. 141.
- Fehr, Ernst, 1986, "A Theory of Voluntary Equilibrium Unemployment," *Journal of Institutional and Theoretical Economics* 142, pp. 405-430.
- Fehr, Ernst, Gächter, Simon, and Kirchsteiger, Georg, 1996, Reciprocal Fairness and Noncompensating Wage Differentials," *Journal of Institutional and Theoretical Economics* 152:4, pp 608-640.
- Fehr, Ernst, and Falk, Armin, 1999, "Wage Rigidity in a Competitive Incomplete Contract Market," *Journal of Political Economy* 107, pp 106-134.
- Fehr, Ernst, and Schmidt, Klaus, 1999, "A Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics* 114, pp 817-868.
- Fehr, Ernst, Gächter, Simon, 2000, "Cooperation and Punishment in Public Goods Experiments," *American Economic Review* 90(4), pp 980-994.
- Fehr, Ernst, Klein, Alexander, and Schmidt, Klaus, 2000, "Endogenous Incomplete Contracts," mimeo, University of Munich.
- Fehr, Ernst, and Gächter, Simon, 2000, "Do Incentive Contracts Crowd out Voluntary Cooperation?" Working Paper #34, Institute for Empirical Research in Economics, University of Zurich.
- Fehr, Ernst, Gächter, Simon, 2005, "Wage Differentials in Experimental Efficiency Wage Markets," in *Handbook of Experimental Economics Results*, Plott, Charles, and Smith, Vernon, editors, North-Holland Press, Amsterdam.
- Güth, Werner, Schmittberger, Rolf, and Schwarze, Berndt, "An Experimental Analysis of Ultimatum Bargaining," *Journal of Economic Behavior and Organization* 3(4), pp. 367-388.
- Hamilton, Barton, Nickerson, Jack, and Owan, Hideo, 2003, "Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation," *Journal of Political Economy* 111(3), pp. 465-497.
- Krueger, Alan, and Summers, Lawrence, 1988, "Efficiency Wages and the Inter-industry Wage Structure," *Econometrica* 56, pp. 259-294.
- Leibenstein, Harvey, 1987, *Inside the Firm: The Inefficiencies of Hierarchy*, Harvard University Press, Cambridge, Massachusetts.
- Martin, I.T., 1915, "The Melting Pot at Ford: Conditions of Employees under the new Profit-Sharing System," *Official Journal of the Carriage, Wagon, and Automobile Workers Union* 4 (August), pp. 8.
- MacLeod, Bentley, and Malcomson, James, 1989, "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment," *Econometrica* 57, pp. 447-480.
- Pliny The Elder, *Natural History: A Selection*, Healey, John, ed., Penguin Classics, New York 1991.
- Rabin, Mathew, 1993, "Incorporating Fairness into Game Theory and Economics," *American Economic Review* 83:5, pp. 1,281-1,302.
- Raff, Daniel M.G., 1988, "Wage Determination Theory and the Five Dollar Day at Ford," *Journal of Economic History* 48(2), pp 387-399.

Shapiro, Carl, and Stiglitz, Joseph, 1984, "Equilibrium Unemployment as a Worker Discipline Device," *American Economic Review* 74, pp. 433-44.