The Gift of Advice: Communication in a Bilateral Gift Exchange Game

3/16/2011

David J. Cooper
Department of Economics
Florida State University
Tallahassee, FL 32306-2180

John P. Lightle
Department of Economics
Florida State University
Tallahassee, FL 32306-2180

Abstract: Studies of gift exchange have focused on employees' provision of effort to reciprocate good wages, but other avenues of reciprocation exist. We study a natural alternative, the use of verbal rewards and punishments, by altering a bilateral gift exchange game so that employees can send messages as well as setting an effort level after observing the wage set by their employer. A simple model of fairness, reciprocity, and communication, similar to the model of Cox, Friedman, and Gjerstad (2007), predicts that allowing messages should lead to lower effort, controlling for wages, and higher wages. In reality effort is virtually unaffected by allowing messages but wages are dramatically increased. This effect occurs because employees give their employers advice to set higher wages. Since the profit maximizing wage is relatively high, employers who follow such advice do well and learn to use higher wages.

Key Words: Gift exchange, communication, advice, experiment

JEL: C70, C92, D23, J30

Acknowledgments: We would like to thank Mark Owens and John Kagel for sharing their experimental software with us. John Jensenius and David Johnson provided us with valuable research assistance. We would like to thank Jordi Brandts, Dan Friedman, Dan Houser, John Kagel, Quoc Tran, and seminar participants at Florida State University and the Southern Economic Association meetings for providing useful advice.

1. Introduction: Motivated by Akerlof's (1982) theoretical work on gift exchange, path-breaking experimental work by Fehr, Kirchsteiger, and Riedl (1993) established the existence of a strong positive relationship between wages and effort in labor markets with incomplete contracts.[1] This finding has been replicated in numerous studies.[2] These experiments share a common feature: reciprocity is captured through a single channel, the effort expended by workers. In lab experiments effort is usually the sole channel employees are given to respond to the wage set by their employer. Other channels might exist in field experiments, but the data analysis typically focuses on how workers' output (or quality of output) changes in response to changes in wages. In reality there are many ways employees can reciprocate the wage choices of their employer. If the existence of alternative means of reciprocation affects the relationship between effort and wages, then experiments neglecting the effect of alternative channels of reciprocation may misrepresent this relationship. The presence or lack of alternative means of reciprocation may also provide a partial explanation for why gift exchange varies across settings and populations.[3] We address these issues by studying a natural alternative channel for reciprocity, one-way communication from employees to employers sent at the same time effort is selected.

Specifically, we present results from a laboratory experiment where subjects assigned the roles of employers and employees play a series of twenty bilateral gift exchange (henceforth BGE) games (Fehr, Kirchler, Weichbold, and Gächter, 1998; Brandts and Charness, 2004; Kagel and Owens, 2010). In each round, employers and employees are randomly matched to play a single BGE game. Play begins with the employer choosing a wage. The employee observes this wage and responds by choosing a costly effort level. As a treatment variable the employee also has the opportunity to enter a text message at the same time he chooses an effort level. This is shown to the employer when he observes the employee's effort. All sessions included ten rounds where messages were available and ten rounds where they were not, with the order of the treatments systematically varied.

If subjects are self-regarding (only concerned with maximizing their monetary payoffs), employees have no incentive to choose more than the minimum effort level since the wage cannot be made contingent on the effort. Anticipating this, employers have no reason to provide more than the minimum wage. These predictions are not affected by the inclusion of messages since these are cheap talk.

[1] Specifically, workers cannot pre-commit to an effort level and employers have no means of sanctioning low effort either through enforceable incentive contracts or termination of the employer-employee relationship

[2] Lab studies that find positive gift exchange, specifically a positive relationship between wages and effort, include Fehr and Falk (1999) and Fehr and Gächter (2000) while examples from field experiments include Al-Ubaybli, Andersen, Gneezy, and List (2008), Cohn, Fehr, and Goette, (2009), and Kube, Maréchal and Puppe (2010).

[3] Gneezy and List (2006) observe little gift exchange in a field study and low gift exchange is also observed for undergraduates, but not MBAs, in a lab study by Hannan, Kagel and Moser (2002). We do not claim that low gift exchange in these two studies is due to communication, but instead cite them as evidence that strong gift exchange is not a universal phenomenon.

We did not expect the preceding predictions to be accurate given the ample existing evidence that subjects are not self-regarding. We therefore develop a simple model of fairness and reciprocity, inspired by Cox, Friedman, and Gjerstad (2007),[4] that includes communication by employees. Positive messages are assumed to provide a benefit to the employer while negative messages impose a penalty. This approach is consistent with evidence from public goods games showing that non-pecuniary punishments generate more cooperative behavior (Masclet, Noussair, Tucker, and Villeval, 2003), implying that negative messages act as a punishment even though there are no monetary implications. All messages are assumed to impose a small cost on the employee. The model predicts that effort will be weakly less, holding wages fixed, with employee messages than without. Intuitively, communication is a cheap way of reciprocating the employer's wage choices, so employees substitute away from costly effort in favor of less expensive verbal punishments and rewards. Less intuitively, the *slope* of the effort-wage curve (i.e. effort as a function of wage) is predicted to be greater when communication is available. Wages are predicted to be weakly greater with communication. The model makes no predictions about learning since it assumes that employers know the utility function of employees and hence know how employees will respond to wages. Realistically, it seems unlikely that employers would know the exact relationship between wages and effort in the absence of experience. We therefore anticipate that employers' choices will change as they have a chance to learn about the response of employees to differing wages.

The theory gets limited support from the data. Effort is not lower when communication is possible, controlling for wages, and the slope of the effort-wage curve is not significantly steeper. As predicted, wages are significantly higher when communication is possible. The effect of messages on wages has a strong dynamic element, taking several periods to emerge. It is a persistent effect with wages continuing to be high even after the possibility of sending messages is removed. Because the high wages with messages lead to high effort levels in response, payoffs for both roles increase when messages are (or have previously been) possible. Thus, allowing messages is efficiency enhancing.

The model correctly predicts higher wages with messages, but does a poor job of explaining the reasons for this effect. To understand why messages have a powerful effect on wages, we coded all of the employees' messages for three broad categories of content: positive feedback (verbal rewards), negative feedback (verbal punishments), and advice on how to play the BGE game. Based on the model we expected communication to increase wages through verbal rewards and punishments, but neither has an effect on future wages chosen by the recipient. Instead, advice has a strong positive effect on future wages

[4] We use a different functional form that Cox *et al*. In developing the model, we restricted ourselves to functional forms that generate predictions consistent with the results of Xiao and Houser's (2005) experiment studying ultimatum games with messages by responders. The functional form used by Cox et al leads to a prediction that adding messages will lead to more rejections, exactly the opposite effect from what is observed.

chosen by the recipient. This makes sense when the content of the advice is considered. Most advice tells the employer that he should have chosen a higher wage, often adding that a higher wage would lead to higher effort. Given the matching mechanism, such advice cannot be interpreted as a direct promise of a reward by the employee to the employer. Instead, we conjecture that the effect of advice is due to the information it contains. Advice gives employers a counter-factual they would otherwise lack. Because of the strong relationship between wage and effort in our data, choosing a wage at or near the maximum is always payoff maximizing for employers. However, it is difficult for employers to realize the benefits of setting high wages if they rarely choose a high wage. Without messages, many employers never experiment with setting a high wage. With messages, advice helps employers who consistently set low wages become aware of the potential gains from choosing a high wage even if they haven't directly experienced this outcome by choosing a high wage. Once employers try a high wage and receive a high effort in return, they tend to stick with high wages.

To summarize, allowing employee messages changes the nature of gift exchange, but not because it provides a second channel for reciprocation. We incorrectly assumed that the effect of messages would depend on the use of verbal rewards and punishments. Instead, allowing messages affects wages because of advice from employees to employers. The existing literature pay little attention to the possibility that gift exchange must be learned, but even if employers initially understand that there is a positive relationship between wages and effort, it stretches to credulity to think they would know the exact strength of that relationship. Messages, specifically advice, give employers information they would not otherwise have and make it easier for them to learn the benefits of gift exchange.

The existing paper that most closely resembles our work is Xiao and Houser's (2005) study of ultimatum games where responders can send a message at the same time they accept or reject an offer.[5] They find that rejection rates are lower, controlling for offers, when messages are possible. Given that the main result in our work is dynamic in nature, there is no possibility it could have been observed in Xiao and Houser's data since they use a one-shot design. We do find it puzzling that they found an effect on rejection rates while we observe no effect on wages. Further work is needed on how verbal rewards and punishments interact with the use of pecuniary rewards and punishments.[6]

Our results provide a potential explanation for why gift exchange varies across settings, but not the one we anticipated. Allowing employees to send messages does not change the relationship between

---

[5] See Xiao and Houser (2009) for related results with dictator games.

[6] In other related work, Noussair and Tucker (2005) study public goods games in which monetary and non-pecuniary punishment are available. Adding non-pecuniary punishment leads to a slight increase in contributions and a more substantial increase in earnings due to lower use of monetary punishments. Non-pecuniary punishment is limited to the assignment of punishment points, so the effects we observe from advice are not possible.

effort and wages in our experiment, but instead makes it easier for employers to learn about the benefits of setting high wages. This suggests that firms are more likely to take advantage of gift exchange when it is relatively easy for them learn about employees' likely responses to increased wages, possibly through advice from employees as in our experiments. It also suggests that advice from employees will only have an effect when gift exchange is sufficiently strong that setting high wages is profitable. Employees presumably always want to tell their employers that wages ought to be higher. This has a strong effect in our experiment because high wages are mutually beneficial.

The remainder of this paper is organized as follows. Section 2 describes the BGE and lays out the details of the experimental design. Section 3 develops theoretical predictions about the effects of adding employee messages to the BGE. The experimental results, including analysis of the content of messages, are described in Section 4. Section 5 discusses the results.

2. Experimental Methods

A. The BGE Game: The BGE game is a sequential game played between two individuals, an employer and an employee. It starts with the employer choosing a wage w for the employee. The employee observes w and chooses an effort level e. The available values of w and e are restricted to integers in the interval [0,100]. Payoffs are given by equations 1 and 2, where Π_M and Π_E represent the employer and employee payoffs respectively.

$$\Pi_M = 100 - w + 5e \quad \textbf{(1)}$$

$$\Pi_E = 100 - e + 5w \quad \textbf{(2)}$$

These payoff functions, borrowed from Kagel and Owens (2010) who in turn modified the payoff functions of Brandts and Charness (2004), have several nice features beyond simplicity. The payoffs are linear and symmetric, so the marginal benefit to your partner (and cost to yourself) of giving a gift does not depend on your partner's actions and is the same for both an employer providing wages and an employee providing effort. There is a unique surplus maximizing outcome with the employer choosing the maximum possible wage (w = 100) and the employee choosing the maximum possible effort (e = 100). This yields equal payoffs to the two players so equity and efficiency are not in conflict. Our use of a relatively high efficiency factor – an increase in wages (effort) by one unit increases payoffs for the employee (employer) by five units – is intended to make effort relatively responsive to wages and make setting high wages relatively profitable for an employer. With a lower efficiency factor we would expect weaker gift exchange and less impact on wages from allowing employee messages.

B. Treatments and Procedures: Our experimental design contains two treatments: *Messages* and *No Messages.* The only difference between these treatments is that in the *Messages* treatment employees can type a message to their employer at the same time they choose an effort level. We use a within-subject design, so all subjects participate in both treatments but the order in which the treatments occurred varied by session.

We conducted six sessions at the xs/fs laboratory at Florida State University, each containing between 22 and 26 subjects. Subjects were recruited with the software ORSEE (Greiner, 2003) and participated in the experiment via a computer network using z-Tree (Fischbacher, 1999). In each session the subjects were randomly divided into two groups of equal size, called “Employers” and “Employees,” and told that they would remain in this role throughout the session.[7] There were 20 rounds per session. Subjects were paid based on their cumulative earnings from all 20 rounds, converted at a rate of 400 experimental dollars equal 1 dollar, plus a $10 show-up fee. The sessions lasted approximately 1¼ hours and the average earnings were $26.23.

Sessions began with the subjects receiving instructions (see Appendix A for a copy). The instructions were read aloud while the subjects read along and were permitted to ask questions. After the instructions were read, the subject took a brief payoff quiz. Round 1 of the experiment did not begin until all subjects had correctly answered the payoff quiz.

In three sessions, the first 10 rounds were the *Message* treatment and the last 10 rounds were the *No Message* treatment. In the other three sessions, the treatments were conducted in reverse order. We refer to *MessagesFirst* and *MessagesSecond* sessions to distinguish the orderings. Subjects knew they would be receiving new instructions prior to Round 11, but were not told any details about the upcoming changes to prevent them from altering their behavior in anticipation.

A zipper or absolute stranger matching was used so that each employee was matched with a different employer in every round during a 10-round treatment. A subject could be re-matched with the same person at most once, and the second encounter between two individuals could only occur in the second 10-round treatment.[8] The instructions for the experiment stressed that subjects could not be matched more than once within a ten round block. There were no means of identification for either type, and no way of knowing if or when a repeated interaction would occur. The purpose of this matching procedure

[7] Use of meaningful context has become standard in gift exchange experiments as a way of helping subjects to understand the structure of the game.
[8] Using an absolute stranger matching over all 20 rounds would require more subjects than the xs/fs lab can hold.

is to make the experiment a series of one-shot games in the minds of the subjects, limiting any reputational effects.

In the *Message* treatment, employees were given 30 seconds after seeing the wage to type messages which would appear in real time on the computer screens of their employers.[9] The employees were told they were "free to type" messages, but were not explicitly told to do so. Employees were given no guidance about the content of their messages other than to avoid profane, threatening, or self-identifying language.[10] Employees could enter a non-binding effort level at any time, but their final decision for effort was not made until after the 30 second message period had elapsed. All messages are after-the-fact in the sense that they are composed after the wage has been determined and at essentially the same time as effort is chosen. Messages therefore cannot be directly responsible for any tacit coordination or implicit contract between employer and employee to achieve high wages and high efforts. In the *No Message* treatment, the employees did not have the ability to send a message, and instead were instructed to enter an effort level after observing the wage. After the employee's decision, both players observed the wage, effort, and payoffs for that round.

3. Theoretical Predictions

If subjects are self-regarding (solely concerned with maximizing own payoffs), the unique subgame-perfect equilibrium of the BGE game is for the employee to choose an effort of 0 regardless of wage, and the manager to assign a wage of 0 in anticipation of this fact. Adding messages by employees has no effect on this prediction since the messages are cheap talk. However, the well-documented positive relationship between wages and effort makes it clear that predicting the effect of messages requires a model where subjects are not self-regarding. In designing a model, we used several criteria.

1) The positive relationship between wages and effort in the BGE can be captured by models with purely distributional preferences. However, the results of Charness (2004) indicate that employees' provision of costly effort is not solely based on distributional concerns but also contain an element of reciprocity. We therefore use a model that incorporates reciprocity. Reciprocity plays little role in our main theoretical findings, but plays an important role in our analysis of alternative models of messages' effects.
2) Models such as Dufwenberg and Kirchsteiger (2004) that draw on psychological game theory and model reciprocity as part of an equilibrium are extremely elegant, making reciprocity endogenous

[9] Thirty second is long enough for employees to send multiple messages to their employers, but not so long that the session grinds to a halt if one employer is especially loquacious. Our impression is that employees were not restricted from sending messages by the thirty second time limit.
[10] There are no cases where employees identified themselves in their messages.

rather than imposing it exogenously. However, these models are hard to work with because they require us not only to be concerned with actions but also to model first and second order beliefs. They often yield multiple equilibria and therefore may not lead to clean predictions. We want a model that is tractable and makes clear predictions. We therefore use a model with that imposes a relationship between actions and perceived kindness rather than deriving kindness as part of an equilibrium. This requires an assumption that perceived kindness is an increasing function of wages. In our opinion this assumption is innocuous.

3) In predicting results for our new experiments we also want to be consistent with the results of existing experiments. Xiao and Houser (2005) find that rejection rates in the ultimatum game are lower, controlling for offers, when responders can send messages. Having responders send messages in an ultimatum game is closely related to having employees send messages in the BGE game. In both cases a sequential game is played where there the second player responds to the first player's action (offer or wage) by either taking a costly action (rejection or effort) or sending a message. If we claim that our theory will predict the effects of employee messages in the BGE game, it should also predict that responder messages will lead to lower rejection rates in the ultimatum game.

In designing a model, our starting point was Cox, Friedman, and Gjerstad's (2007) model of fairness and reciprocity. Cox *et al* model an individual's concern for the payoffs of others as depending on a single parameter, their emotional state (θ). A second mover is willing to give up some of his payoff to increase the first mover's payoff when he is in a positive emotional state ($\theta > 0$) and would like to decrease the first mover's payoff if he has a negative emotional state ($\theta < 0$). Positive and negative reciprocation can be captured by these preferences because the second mover's emotional state, θ, is an increasing function of the benevolence of the first mover's actions and can change sign accordingly. Cox *et al* provide a tractable model of other-regarding preferences that incorporate reciprocity.

Employee messages allow for the possibility of non-pecuniary reciprocation for high wages. A natural way to incorporate this into Cox *et al*'s model would be to assume that receiving a positive message increases the employer's payoff a given amount while receiving a negative message decreases his payoff. However, the modified model leads to a counterintuitive prediction. When an individual is in a negative emotional state ($\theta < 0$), the marginal rate of substitution between his own payoff and the other player's payoff decreases in absolute value as the other's payoff decreases. In other words, holding the emotional state fixed, an individual's willingness to punish *increases* as the other player's payoff decreases. This implies that rejection rates, controlling for offers, should be *higher* in the ultimatum game when responders can send messages, rather than *lower* as observed by Xiao and Houser (2005).

This incorrect prediction is not a general property of models that include an emotional state, like Cox *et al*, but instead follows from the specific functional form used by Cox *et al*.

We therefore construct a model in which, keeping the emotional state constant, individuals are more willing to reward others as the other's payoff decreases, and less willing to punish others as the other's payoff decreases. This is accomplished by letting an individual's willingness to exchange their payoff for other's payoffs depend on whether or not the exchange is in the domain of punishments or rewards. Consider a utility function defined over own payoff (m), rewards to the other player (r), and punishments of the other player (p). Let y be the other player's payoff, let $\underline{y}$ be the minimum possible monetary payoff for the other player, and let $\overline{y}$ be the maximum possible monetary payoff for the other player. We model an individuals' utility as having the following functional form:

$$u(m,r,p) = m^{(1-\theta_r-\theta_p)} r^{\theta_r} p^{\theta_p} \qquad \textbf{(3)}$$

$$r \equiv \max\left[y - \underline{y}, 0\right] \qquad p \equiv \max\left[\overline{y} - y, 0\right] \qquad \textbf{(4)}$$

The player's emotional state is captured by the parameters θ_r and θ_p. We assume $0 \leq \theta_r \leq 1$ and $0 \leq \theta_p \leq 1$. The parameters θ_r and θ_p are assumed to be mutually exclusive in the sense that if $\theta_r > 0$ then $\theta_p = 0$ and vice versa. Intuitively, if $\theta_r > 0$ then the player is in a positive emotional state and wishes to reward the other player. If $\theta_p > 0$ then the player is in a negative emotional state and wishes to punish the other player. A player cannot wish to reward and punish the other player simultaneously. We refer to θ_r and θ_p as the "reward state function" and "punishment state function". Holding the emotional state (θ_r and θ_p) fixed, the resulting utility function is a Cobb-Douglas utility function over two goods. A Cobb-Douglas utility function was chosen for the sake of tractability, but any well-behaved homothetic utility function would yield similar results.

Consider the ultimatum game where a proposer demands an amount D from a fixed pie P and a responder chooses to accept or reject it. If the demand is accepted, the proposer gets D and the responder gets P – D. Otherwise both players get zero. On top of their payoffs from the game, both players get a fixed show-up fee of $s > 0$. Applying our model, let θ_p and θ_r be the punishment and reward state functions for the responder. These are assumed to be continuous functions of D as well as differentiable when the functions are strictly positive. Assume that $\theta_p(D) > 0$ iff D > P/2 and $\theta_p'(D) > 0$ if D > P/2. It is easily proved that there exists D* such that P/2 < D* < P and all demands D > D* are rejected.

Now suppose responders can send a message back to the proposers at the same time the demand is accepted or rejected. We model a negative message as imposing a "psychic" cost of B on the proposer while a positive message provides a "psychic" benefit of B. The responder's cost of sending the message is C, which for simplicity is assumed to be arbitrarily close to zero. Costs and benefits due to messages are modeled as additions or subtractions from the players' payoffs, with both players' utilities being affected by messages solely through these changes in payoffs.[11] Assume $\theta_p(D)$ is independent of either the possibility or content of messages.

Claim 1: *Any responder who is indifferent between accepting and rejecting a demand if messages are not available will strictly prefer to accept the same demand if messages are available.*

Proofs for all claims, propositions, and corollaries are located in Appendix B. Intuitively, a responder will certainly punish a proposer who makes an unsatisfactory offer by sending a negative message. Given that this lowers the proposer's payoff, it also reduces the responder's willingness to punish in the form of a rejection. As a result, responders are more likely to accept an offer at the margin when messages are available. It follows from Claim 1 and monotonicity of the utility function that a responder will accept any offer when messages are available that he would have accepted without messages. Therefore, rejection rates are weakly lower when messages are available, consistent with the results of Xiao and Houser (2005).

Now consider the implications of our model for the BGE game with messages. Once again assume that both players receive a fixed show-up fee of $s > 0$ on top of their payoffs from the game. As in the ultimatum game, messages are assumed to affect payoffs for employers and employees additively. A negative message imposes a "psychic" cost of B to the employer while a positive message gives a "psychic" benefit of B. The employee's cost of sending the message is C, which again is assumed to be arbitrarily close to zero. If a positive message is sent, the players' payoffs from (1) and (2) are adjusted to (5) and (6) to incorporate the benefit and cost of the message as well as the show-up fee.

$$\Pi_M = 5e - w + 100 + B + s \quad \textbf{(5)}$$

$$\Pi_E = 5w - e + 100 - C + s \quad \textbf{(6)}$$

If a negative message is sent, (5) is replaced with (5').

$$\Pi_M = 5e - w + 100 - B + s \quad \textbf{(5)}$$

The emotional state of employers is assumed to be neutral ($\theta_r = \theta_p = 0$). Intuitively, an employee hasn't taken any action yet when their employer chooses a wage, and therefore cannot be perceived as

[11] For example, suppose a proposer makes a demand of D, has this demand accepted, but also receives a negative message. The proposers payoffs is D + s – B and the responder's payoff is (P – D) + s – C.

having acted in a kind or unkind fashion. For employees, θ_p and θ_r are functions of the wage w. We again assume continuity and differentiability (if the functions are strictly positive). We assume $\theta_r(0) = 0$, $\theta_p(100) = 0$, $\theta_r'(w) \geq 0$ with strict inequality if $\theta_r(w) > 0$, and $\theta_p'(w) \leq 0$ with strict inequality if $\theta_p(w) > 0$. An employee's emotional state is assumed to be independent of the possibility and content of messages.

Let $e^*(w, \theta_r(w))$ be the utility maximizing effort when messages are not available. This is a function of the wage w directly through its first argument and indirectly through the reward state function $\theta_r(w)$ in the second argument. Let $e^{**}(w, \theta_r(w))$ be the analogous utility maximizing effort when messages are available. Notice that the punishment state function, $\theta_p(w)$, does not appear as an argument in these functions. The utility maximizing effort is always zero if $\theta_r = 0$, so the magnitude of $\theta_p(w)$ does not affect the optimal effort.

Proposition 2 establishes that the availability of messages leads an employee to choose an effort which is less than or equal to the effort he would have chosen without messages. This is true for any wage received, and for any reward state function. Corollary 3 shows that the effort is *strictly* lower when messages are available for (almost) all cases where employees are not in a corner solution when messages are not available.

Proposition 2: *Given any wage and reward state function,* $e^*(w, \theta_r(w)) \geq e^{**}(w, \theta_r(w))$.

Corollary 3: *For a given wage and reward state function,* $e^*\big(w, \theta_r(w)\big) > e^{**}(w, \theta_r(w))$ *if* $B > 5C$ *and* $100 > e^*\big(w, \theta_r(w)\big) > C$.

Our model makes an unambiguous prediction that wages will be lower when messages are possible. The intuition behind this result is straightforward. Providing effort is one way for employees to increase the manager's payoff at a rate of 5 units for every 1 unit given up, but when $B > 5C$, sending positive messages is a cheaper way to increase the manager's payoff. Since we assume the cost of messages is very small, this condition is expected to hold. Therefore, utility-maximizing employees will substitute away from using the relatively expensive method of choosing a high effort to reward the manager and combine the positive message with a reduced level of effort.

We now consider the effect of messages on the shape of the effort-wage curve. Define $\underline{w}$ as the largest wage below which an employee exerts no effort when messages are not available: $\underline{w} = inf\big\{w | e^*\big(w, \theta_r(w)\big) > 0\big\}$. Similarly, let $\underline{w}'$ be the largest wage below which an employee exerts no effort when messages are available: $\underline{w}' = inf\big\{w | e^{**}\big(w, \theta_r(w)\big) > 0\big\}$. We refer to $\underline{w}$ and $\underline{w}'$ as the "cutoff wage" without and with messages respectively. Let $\left(\underline{m}, \underline{y}\right)$ be the manager and employee payoffs when the wage is $\underline{w}$ and no message is sent. We now make a natural assumption about employees' reward state functions and utility functions.

Assumption 1: *The employee's utility function and reward state function are such that he weakly prefers sending a positive message to sending no message given a wage of* $\underline{w}$: $u\left(\underline{m}, \underline{y}\right) \leq u\left(\underline{m} - C, \underline{y} + B\right)$.

The only costs of sending a message in our experiment are the time and effort it takes to type a message into a computer. We believe these costs are sufficiently low to justify Assumption 1, which states that the benefits of sending a message outweigh the costs for an employee who has received a wage which puts him right on the margin of returning zero or positive effort. Given Assumption 1 we can prove the following two propositions about how the shape of an employee's effort-wage curve changes if employee messages are allowed. Proposition 4 states that's the employee's cutoff wage without messages is strictly smaller than the employee's cutoff wage with messages as long as there is some wage that leads to positive effort when messages are not possible.

Proposition 4: *If* $\underline{w} < 100$, *then* $\underline{w} < \underline{w}'$.

The intuition behind Proposition 4 is that for some wage slightly above the cutoff wage without messages, the employee would choose to send positive effort when messages are not available, but would instead send a positive message and zero effort if messages were available. This is because, given Assumption 1, messages are a relatively cheap way to increase the manager's payoff. Proposition 5 addresses the slope of the effort-wage curve.

Proposition 5: *Given Assumption 1, the slope of the effort-wage curve when messages are available is strictly steeper than when messages are not available for any w such that* $0 < e^{**}(w, \theta_r(w)) < 100$.

The intuition underlying Proposition 5 is as follows. For any wage where a positive message is sent, the employee's optimal choice leads to higher payoffs (m and y) for both players than in the absence of messages. With or without messages, the first-order condition for utility maximization at an interior solution is $\left(\frac{1-\theta_r}{\theta_r}\right)\frac{y}{m} = 5$. As θ_r increases, effort must rise in order to maintain this equality by increasing the ratio (*y*/*m*). Since *y* and *m* are both larger with messages and the ratio (*y*/*m*) is the same with or without messages, the change in effort required to produce the needed change in (*y*/*m*) is also higher with messages. Therefore the effort-wage curve must be steeper when messages are available.

Propositions 2, 4, and 5 jointly characterize how the effort-wage curve changes if employee messages are allowed. Figure 1 illustrates these changes for a specific parameterization of the model: B = 100, C = 1, $\theta_r(w) = (w^2/20{,}000)$. Starting at a wage of zero, optimal effort is zero with or without messages. As wages increase, the effort-wage curve remains at zero for longer when messages are possible. Once both effort-wage curves start to increase, the effort chosen with messages remains less than the effort chosen without messages. However, the effort-wage curve when messages are available is steeper over the range of wages where interior choices of effort are made. Allowing messages has a more dramatic effect on the level of effort (holding wages fixed) than the slope of the effort-wage curve. This is not due to the specific parameterization used to generate Figure 1, but instead reflects a general property of the model. Because the effect of messages on the slope of the effort-wage curve is indirect, occurring through differing impacts from changing θ_r, it is more subtle than the effect on the level of effort.

[Figure 1 here]

We now turn to the effect on wages of allowing messages. Since the emotional state of the employer is assumed to be neutral, utility maximization is equivalent to payoff maximization for the employer. Let w^* be the payoff-maximizing wage when messages are not available, and w^{**} be the payoff-maximizing wage when they are.

To understand the relationship between w* and w**, we first need to characterize the optimal wage when messages are not possible. Let $\hat{w}$ be the lowest wage that elicits the maximum achievable effort from an employee in the absence of messages. If there exists a wage such that $e^*\big(w, \theta_r(w)\big) = 100$, $\hat{w}$ is the lowest wage that induces 100 as the effort. If there is no wage that induces positive effort, $\hat{w} = 0$. Otherwise, $\hat{w} = 100$. The experimental parameters and assumptions about the employee's utility ensure that $\frac{de^*}{dw} > \frac{1}{5}$ if the employee is choosing an interior level of effort. Given that a unit of effort is worth five experimental dollars to the employer, it follows that the employer can always do better by raising the wage if the employee is choosing an interior level of effort. This implies that either $w^* = 0$ or $w^* = \hat{w}$.

We now argue that $w^{**} \geq w^*$ almost always. For any case where $w^* = 0$, this is trivial. Suppose $w^* = \hat{w} > 0$. Let $\hat{w}'$ be the lowest wage that elicits the maximum achievable effort from an employee with messages possible. If $0 < \hat{w} < 100$, then $\hat{w}' > \hat{w}$ by Corollary 3. Let $\tilde{w}$ be the lowest wage that yields a positive message. As long as C is sufficiently small relative to B, effort equals zero at $\tilde{w}$. Because $\frac{de^{**}}{dw} > \frac{1}{5}$ and $w^* \neq 0$, the optimal wage with messages must either be $\tilde{w}$ or $\hat{w}'$. If $0 < \hat{w}' < 100$ then $w^{**} = \hat{w}' > \hat{w} = w^*$ as desired. However, if $\hat{w}' = 100$ then it is possible to generate examples where $w^{**} = \tilde{w} < w^*$. We argue that these cases are very unlikely. For example, consider reward state functions in the form $\theta_r(w) = kw$ where k > 0. For simplicity assume that C is arbitrarily small so the employee sends a positive message for any w > 0. If $\frac{1}{3000} < k < \frac{1}{3000} + \frac{29B}{9{,}000{,}000+3000B}$, then $w^* = 100$ and $w^{**} = \widetilde{w} < w^*$. Obviously this is a small target to hit, but it is a possibility. More generally, the maximum an employer can earn by choosing $\widetilde{w}$ is 100 + B. If he sets a wage of 100, his payoff is B + $5e^{**}\big(100, \theta_r(100)\big)$. Therefore, a sufficient condition for w** = 100 is $e^{**}\big(100, \theta_r(100)\big) > 20$. In our experimental data, the average effort following a wage of 100 consistently hovers in the mid-50s regardless of time period or treatment. We therefore feel comfortable predicting that wages will be weakly higher with messages.

We summarize the predictions of our model with the following three hypotheses.

Hypothesis 1: Holding wage fixed, effort will be lower when messages are possible.

Hypothesis 2: The slope of the effort-wage curve will be steeper when messages are possible.

Hypothesis 3: Wages will be higher when messages are possible.

These hypotheses are based on the predictions of a specific model that incorporates specific assumptions about subjects' preferences and a specific functional form. It is reasonable to ask how much our hypotheses depend on our choice of models. Hypothesis 1 turns out to be extremely robust. Consider three alternative models: The Cox *et al* model of fairness and reciprocity, Bolton and Ockenfels' model of inequality aversion (Bolton and Ockenfels, 2000),[12] and our model with "emotional dampening". The final alternative requires some explanation. Xiao and Houser (2005) conjecture that messages reduce rejection rates in the ultimatum game by allowing responders an alternative means of expressing their emotions. This implies that the presence of messages dampens the emotional state of a responder in the ultimatum game, or the employee in a BGE game. A simple way to model this is to scale down the emotional state (reward or punishment) by a factor of k when messages are possible, where 0 < k < 1. Because messages let employees vent their emotions, they are inherently less kind when responding to a high wage. This is the only effect of messages in the model with emotional dampening. In other words, B = C = 0. All three of these models predict that holding wages fixed, effort will be lower when messages are possible. Thus, we are most confident in Hypothesis 1. Hypotheses 2 and 3 are less robust. The model with emotional dampening, for example, predicts that the slope of the effort-wage curve will

[12] We picked Bolton and Ockenfels' model as a very well known example of distributional preferences. Bolton and Ockenfels' model is slightly easier to work with for our purposes than Fehr and Schmidt's (1999) model because it typically gives interior solutions, but this is not a general statement about the relative merits of the models.

be flatter with messages and it is straight forward to derive examples where wages are lower with messages.

Our model is static. However, it is highly likely that employees are heterogeneous, with different types having different reward state functions and, by extension, different choices of effort as a function of the wage. If the distribution of types is unknown to employers, a reasonable assumption, the theory can naturally be extended to allow for dynamic behavior on the part of employers as they learn about the distribution of types of employees. Without knowing the initial beliefs of employers we cannot predict a direction that wages will adjust, but we do predict that wages will be a function of past effort received as well as past messages.

4. Results

[Table 1 Here]

A. Treatment Effects: Table 1 summarizes wages, efforts, employer payoffs, and employee payoffs broken down by session type (*Messages First* or *Messages Second*) and time period (Rounds 1 – 10 or Rounds 11 – 20). Table 1 also provides t-tests to compare the mean value across treatments, holding the time period fixed, or across time periods holding the treatment fixed. An observation for these tests is the average value of the variable (wage, effort, etc.) for an individual subject across the ten round time period. In other words, each observation corresponds to a single subject. The t-tests reported below that compare across time periods, holding the session type fixed, account for observations being paired. The null hypothesis for all the t-tests is that the means are equal across cells.

For Rounds 1 – 10, wages are significantly higher with messages than without. Allowing messages also leads to higher wages in Rounds 11 – 20 of the *Messages Second* sessions than in Rounds 1 – 10 of this session type. The effect of allowing messages on wages is persistent as wages are significantly higher in Rounds 11 – 20 of the *Messages First* sessions, when messages are no longer possible, than in either Rounds 1 – 10 of this session type, when messages are possible, or in Rounds 1 – 10 of the *Messages Second* sessions, where messages are not possible and subjects have not previously experienced messages.[13] Wages in Rounds 1 – 10 of the Messages First treatment are higher than in Rounds 11 – 20 of Messages Second treatment, but the difference is small and not statistically significant.[14] There is little persistent effect from an initial inability to send messages.

[13] Comparing Rounds 1 – 10 of the *Messages Second* sessions with Rounds 11 – 20 of the *Messages First* sessions, the t-statistic equals 4.888 which is significant at the 1% level.
[14] The t-statistic equals .484.

The treatment effects on effort parallel those for wages, but are statistically weaker. Effort is higher in Rounds 1 – 10 with messages than without, but this difference is significant at the 5% level rather than the 1% level. Effort is also higher in Rounds 11 – 20 of the *Messages Second* sessions than in Rounds 1 – 10 of this session type, but the difference again is significant at the 5% level instead of the 1% level. The effect of allowing messages leads to persistently higher effort as can be seen by comparing Rounds 11 – 20 of the *Messages First* sessions with Rounds 1 – 10 of the *Messages Second* sessions. Again, the difference is statistically significant at the 5% level. It is clear that allowing messages is associated with higher effort, but it remains to be seen whether this is a direct effect of the treatment or is due to the higher wages associated with allowing messages.

Given the high efficiency factors associated with wages and effort, we would expect the increased levels of wages and effort associated with allowing messages to translate into higher payoffs for both employers and employees. This is indeed the case as can be seen in the lower half of Table 1. In keeping with the stronger effect of allowing messages on wages than on effort, the absolute effect of allowing messages on payoffs is larger for employees than for employers, as well as being more statistically significant. As implied by the persistent effects of allowing messages on wages and effort, there are also persistent effects on payoffs for both roles.

Increasing payoffs for both employers and employees implies that total surplus must also be higher when messages are possible. For Rounds 1 – 10, 61% of the possible gains through gift exchange are realized when messages are possible, as opposed to only 43% when messages are not possible. [15] Efficiency rises in both session types for Rounds 11 – 20, reaching 66% of the possible gain in the *Message First* sessions and rising to 55% in the *Messages Second* sessions.

Figure 2 illustrates how wages, effort, and total surplus develop over time in the *Messages First* and *Messages Second* sessions. The effects associated with allowing messages take time to develop. In Round 1, the increases associated with allowing messages are tiny for wages, effort, and total surplus. None of these differences are statistically significant at even the 10% level.[16] The difference between treatments steadily increases for all of these variables over Rounds 1 – 10. Along similar lines, there is little difference in wages, effort, and total surplus between Round 1 and Round 11 of the *Messages*

[15] The total surplus is 200 in the Nash equilibrium and 1000 if total surplus is maximized, so the maximum possible gain through gift exchange is 800 experimental dollars. The percentage of possible gains through gift exchange realized is (total surplus – 200)/800.

[16] This conclusion is based on t-tests using Round 1 data. The t-statistics for wages, effort, and total surplus are 0.772, 0.035, and 0.505 respectively.

Second sessions.[17] Over time, the wages, effort, and total surplus grow once employee messages are allowed. The effect of allowing messages has a strong dynamic component, suggesting that learning of some sort plays a role.

[Figure 2 Here]

The effect on wages of having allowed messages does not die out over time when messages are no long available. Wages in Rounds 11 – 20 of the *Messages First* sessions have no obvious trend either higher or lower. There is a decline in effort over the last few rounds of the *Messages First* sessions, but regression analysis indicates that this decrease is not statistically significant except (possibly) in the final round.[18] The positive effect of allowing messages on wages and effort holds steady over the last few rounds of the *Messages Second* sessions. If wages and effort systematically collapsed over the final few rounds of the experiment, it would suggest that the high levels of gift exchange when messages are (or have been) possible are driven by uncontrolled repeated game or reputation effects. There is little evidence for this.

Conclusion 1: Allowing employees to send messages is associated with higher wages, effort, payoffs for both roles, and total surplus. This effect persists after the possibility of messages is removed.

We have documented that effort is higher when messages by employees are currently allowed or have previously been allowed, but the more relevant question is whether effort is higher *subject to wages*. Figure 3 addresses this question. Wages have been broken down into five categories. In more than half of the observations (780 of 1440), the maximum possible wage of 100 was chosen. These observations form a single category. The remaining 760 observations are broken into quartiles by wage to form the remaining four categories. For each category Figure 3 graphs the average effort conditioned on whether or not messages were possible (*No Messages vs. Messages)*.

[Figure 3 Here]

Regardless of whether or not messages are possible, there is a strong positive relationship between wages and effort. The combination of a high efficiency factor and the strong response of effort to wages makes choosing high wages optimal for employers in both types of sessions regardless of whether or not messages are available. This is not entirely monotonic – if the data is broken down into smaller cells than

[17] Comparing wages, effort, and total surplus for Rounds 1 and 11, the respective t-statistics are 1.331, 0.551, and 0.780. In no case is the null hypothesis that the two means are equal rejected at even the 10% level.
[18] See p. 19.

those shown in Figure 2, employer payoffs are slightly higher from choosing a wage in the 90s than the maximum possible wage of 100.[19]

Contrary to the strongest of our theoretical predictions, effort is not always lower, holding wage fixed, when messages are possible. For high wages, effort is greater when messages are possible, although in no case is the difference between *Messages* and *No Messages* large. The effort-wage curve is steeper when messages are possible, as predicted, but the difference is again slight. From the point of view of an employer, the monetary incentives to offer a high wage are about the same whether or not messages are possible. The data provides little support for Hypotheses 1 and 2 and casts doubt on the mechanism underlying Hypothesis 3.

While allowing messages has little impact on the relationship between wages and effort, experience does affect this relationship. Figure 4 graphs the average effort broken down by time period (Rounds 1 – 10 vs. Rounds 11 – 20) and wage category. Observations with and without messages possible are pooled. For all wage categories the average effort is lower in Rounds 11 – 20 than in Rounds 1 – 10. This parallels the results of Brandts and Charness (2004), who find that effort decreases with experience, controlling for wages.[20]

[Figure 4 Here]

An obvious explanation for effort decreasing with experience is reputation effects, but it would be surprising if this was the case given the use of large sessions and an absolute stranger matching within ten round blocks. Indeed, a closer reading of the data indicates that reputation effects are not the likely cause of decreasing effort. Choice of a wage of 100 is common in both Rounds 1 – 10 (41% of observations) and Rounds 11 – 20 (67% of observations), so there is no lack of data to observe changes for this specific wage. Employees consistently spend substantial amounts on effort when the wage is 100, giving them plenty of room to save on investments in reputation. If the decline of effort was due to the declining value of maintaining a reputation for reciprocation, we'd expect average effort *subject to the wage equaling 100* to also decline. A decrease is observed, but the magnitude is tiny and, as shown in the regression analysis below, statistically insignificant. The changes in the effort-wage curve are more consistent with development of a social norm than reputation effects. Choice of a wage of 100 becomes increasingly common over time and employees always respond to this with a relatively high average

[19] The average employer payoff from choosing a wage between 90 and 99 is 284 averaging across all treatments as opposed to an average of 270 from choosing 100. This difference is not statistically significant, but employer payoffs are significantly higher for either $90 \leq w \leq 99$ or for $w = 100$ than for any lower range of wages.
[20] Charness, Frechette, and Kagel (2004) also find decreasing effort controlling for wages, but this effect is not monotonic since effort increases at first before falling back to its initial levels (see their Table 2).

effort. Departures from the norm of setting a wage equal to 100 lead to increasingly harsh punishments. For the lowest wage category (wage < 15), wages are already so low in Rounds 1 – 10 that little change is possible. The second lowest wage category (15 ≤ wage < 50) does have room for a decrease and, consistent with increasing punishment of departures from the norm, the decrease in effort is particularly steep for this wage category.

The regressions shown in Table 2 put our observations about the relationship between wages and effort on a firmer statistical footing. The dataset for these regressions includes all 1440 observations from 72 employees. Effort is the dependent variable. Because subjects were restricted to choose an effort between 0 and 100, inclusive, the data is censored. A tobit model is therefore used. The standard errors are corrected for clustering at the employee level.

[Table 2 Here]

Independent variables in Model 1 are the wage and dummies for Rounds 11 – 20 and whether messages were possible ("*Messages*"). This regression studies whether the *level* of effort shifted either with experience or with messages being possible controlling for the wage. Consistent with our observations from Figures 3 and 4, there is a significant negative shift in effort for Rounds 11 – 20, but virtually no change in effort *controlling for wages* when messages are possible. In other words, the observed shift in effort when messages are possible can be entirely attributed to higher wages.

Model 2 looks at whether the *effort-wage curve* shifted either with experience or with the possibility of messages. Interactions between wages and the dummies for Rounds 11 - 20 and messages being possible are added as independent variables. Once again the results are consistent with our observations from Figures 3 and 4. The effort-wage curve has a significantly lower intercept and steeper slope for Rounds 11 - 20 than Rounds 1 – 10. The possibility of messages also leads to a lower intercept and steeper slope, but neither effect is statistically significant.[21] By extension, the monetary incentives faced by employers are largely unaffected by allowing messages, eliminating one possible explanation for higher wages when messages are possible.

Model 2 can be modified to focus on how effort following a high wage changes with experience. Define "reverse wage" using the following equation: reverse wage = 100 – wage. Modify Model 2 by replacing wage with reverse wage every place that wage appears, including the interaction terms. With this change the parameter estimate for "Rounds 11 - 20" now measures whether effort changes with experience for a wage of 100. The parameter estimate for "Rounds 11 – 20" equals -4.51 with a standard

[21] The two parameters also fail to be jointly significant (F = 0.71; 2 d.f.; p = 0.49).

error of 7.95. This is not statistically significant at any standard level. The regression results are consistent with our earlier discussion of how effort changes with experience – effort changes little for high wages levels, but instead the effort-wage curve becomes significantly steeper over time. This implies harsher punishment of low wages.

Model 3 extends the results of Model 2 by examining whether the effort-wage curve varies between the four cells of the dataset (Rounds 1 – 10 vs. Rounds 11 – 20 and *No Messages* vs. *Messages*). The base is Rounds 1 – 10 of the *Messages Second* sessions. As independent variables, dummies are included for the other three cells (*No Messages* * Rounds 11 – 20, *Messages* * Rounds 1 – 10, and *Messages* * Rounds 11 – 20). Wage is included as an independent variable along with interactions between wage and the three cell dummies. Model 3 is equivalent to adding two parameters to Model 2 (an interaction between the dummies for *"Messages"* and "Rounds 11 – 20", and an interaction between this term and wages). The improvement to the model's fit from adding these two parameters is not statistically significant, indicating there is little interaction between the effects on the effort-wage curve of having messages possible and experience.[22] The pattern of parameter estimates is what would be expected given the results of Model 2. The effort-wage curve is steepest for *Messages* * Rounds 11 – 20, combining the effects of experience and having messages possible, and flattest for the omitted category of *No Messages* * Rounds 1 – 10 where neither effect is present.

Digressing slightly, we have also used tobit regressions to examine whether the dip in effort observed for the last few rounds of the *Messages First* sessions is statistically significant. In good economist fashion, the answer is a definite maybe. Controlling for wages, we have compared effort in the final round (Round 20) of the *Messages First* sessions with the final round that messages are available in these sessions (Round 10) and the first round that messages are not available (Round 11).[23] The difference is *not* statistically significant in the first case ($t = 1.55$; $p > .10$) but is in the second case ($t = 2.11$; $p < .05$). If we make these comparisons for Round 19 rather than Round 20, neither difference is significant. To limit the impact of random changes from round to round, we have also run equivalent regressions with the data broken into two round chunks. The differences between Rounds 19 – 20 of the *Messages First* sessions and Rounds 9 – 10 or Rounds 11 – 12 of these sessions are not statistically significant. Our best reading of the data is that effort decreases for Round 20 of the *Messages First* sessions, but it is a minor change.

[22] Running a log-likelihood ratio test, $\chi^2 = 1.58$. With two degrees of freedom, this is well below the cutoff of 4.61 for significance at the 10% level.
[23] Data for these regressions are observations from Rounds 10 – 20 of the *Messages First* sessions. The independent variables are the wage and round dummies with Round 20 as the omitted category.

Conclusion 2: Allowing messages does not make effort lower after controlling for wages and does not significantly affect the slope of the effort-wage curve. The data does not support Hypotheses 1 and 2. The effort-wage curve becomes steeper with experience.

B. Content Analysis: Based on Conclusions 1 and 2, employers offer higher wages when messages are possible (or have previously been possible) even though there is no change in monetary incentives. Two explanations for this phenomenon suggest themselves. First, employers' utilities may be affected by the messages they receive even though their monetary incentives are unchanged. The possibility of receiving a kind (unkind) message might be a potential benefit (cost) that affects the wages offered by employers. They may offer higher wages when messages are possible to increase the likelihood that a kind message will be received rather than an unkind message. This is the mechanism underlying the theoretical prediction that wages will be weakly higher with messages: employers will be more willing to offer high wages if this leads to high effort *and* a positive message than if they gain high effort alone. Second, messages may affect learning by employers. When messages are possible wages rise over time, but without messages wages are flat. This suggests that employers are responding to something in the messages by moving to higher (and more profitable) wages. To examine these two possibilities more carefully, we systematically examine the content of messages sent by employees.

We have classified the messages sent by employees into three broad categories:

1) Positive Feedback: These are messages that direct positive sentiments towards the employer. Sometimes these are explicit expressions of thanks, but often the expression of gratitude is implicit rather than explicit. The examples below illustrate the types of messages included in this category. All messages given as examples are transcribed verbatim without any correction of typos, spelling, or grammar.

"thank you!!!! :)"

"You're awesome, smart, and great. I Love You"

2) Negative feedback: These are messages that direct negative sentiments toward the employer. Typically these refer to wage chosen by the employer, as in the following examples:

"bad decision and rude"

"thats messed up u get wat u pay 4 though"

3) Advice: Messages classified as offering advice either give the employer advice about what course of action should be followed in the future (or should have been followed for the current round) or explain that higher wages will lead to higher effort, implying that a higher wage should be chosen. Messages often combined both types of advice, calling for higher wages and explaining the suggested course of action by pointing out that higher wages will lead to higher effort as in the second example below.

"If you use a higher wage people may work more hours"

"WHAT THE HECK?!?! If you gave me 100, I would have given you a 100 and we both would have made 500 way to blow it"

The message categories are *not* mutually exclusive. Many messages include both negative feedback and advice, as in the second example above. Even when no explicitly negative comments are made, advice often implies a negative reaction to the wage just chosen.

Two research assistants independently coded all messages for the presence these three categories. No effort was made to force agreement among coders – the goal was to have two independent codings so errors were uncorrelated. At no point in the coding process was either RA informed about any hypotheses that were to be tested. The RAs were repeatedly told that their job was to capture what had been said rather than why it was said or what effect it had. Coding was binary – a message was coded as a 1 if it was deemed to contain the relevant category of content and zero otherwise. We had no requirement on the number of codes for a message – a coder could check as many or few categories as he deemed appropriate. There was a high degree of agreement between the coders.[24]

[Figure 5 Here]

Figure 5 shows the frequency of each category, averaging across coders, as a function of the round. Only observations where messages are possible are used to calculate these frequencies, so Rounds 1 – 10 reflect observations from the *Messages First* sessions while Rounds 11 – 20 show observations from the *Messages Second* sessions. The time trends in Figure 5 must be interpreted with care since the likelihood of each category is dependent on the wage, and the distribution of wages shifts upwards over time. Restricting the sample to observations with a wage of 100, the probability of positive feedback falls within each ten round block.[25] In contrast, the likelihood of negative feedback rises within ten round

[24] Cohen's κ for the three categories are .92, .71, and .74 respectively.

[25] The frequencies of positive feedback subject to messages being possible and the wage equaling 100 are .61 for Rounds 1 - 5, .43 for Rounds 6 – 10, .58 for Rounds 11 – 15, and .35 for Rounds 16 – 20.

blocks if the sample is restricted to observations with a wage less than 100.[26] Consistent with our observation that experience leads to more punishment of low wages by withholding effort, experience also seems to lead to more punishment of low wages via negative feedback. There is no obvious trend in the frequency of advice if the sample is restricted to observations with a wage less than 100 except for a very steep decrease in the last couple of rounds when advice about future play is somewhat pointless.[27]

The theory developed in Section 3 predicts that a negative message should never be sent if positive effort is chosen. Intuitively, positive effort is driven by a desire to reward kind behavior. Punishing an individual verbally is inconsistent with rewarding them financially. This is not a prediction that relies on the specific functional form we are using – Cox *et al*'s model and models of inequality aversion also predict that negative messages and positive effort are mutually exclusive. The data are mostly in line with this prediction. In the 65 observations where both coders agreed that negative feedback had been sent, a strictly positive effort was chosen in only 13 cases. For three of these observations, the selected effort was 1. We think it fair to interpret this effort choice as an insult, akin to leaving a penny as a tip in an American restaurant. For the remaining ten observations, half of the messages contained advice to give higher wages. The negative comments in these messages can be interpreted as much as criticism for being stupid as criticism for being cheap (e.g. "are you stupid? why would you not give 100? do you not like to make money?"). This is one case where the theory largely gets it right.

[Figure 6 Here]

Figure 6 shows the relationship between wages and the types of messages sent. The same wage categories are used as in Figures 3 and 4. There is a strong relationship between wages and the use of positive and negative feedback – the likelihood of positive feedback is an increasing function of wages while the likelihood of negative feedback is a decreasing function.[28] The relationship between wages and giving advice is less obvious except that advice is virtually never given following a wage of 100.

The strong relationship between wages and the use of positive or negative feedback is consistent with wages being higher when messages are possible because of benefits associated with receiving positive feedback and costs due to receiving negative feedback. However, this cannot be a full

[26] The frequencies of negative feedback subject to messages being possible and the wage being less than 100 are .13 for Rounds 1 - 5, .23 for Rounds 6 – 10, .26 for Rounds 11 – 15, and .41 for Rounds 16 – 20.
[27] The frequencies of advice subject to messages being possible and the wage being less than 100 are .31 for Rounds 1 - 5, .31 for Rounds 6 – 10, .33 for Rounds 11 – 15, and .20 for Rounds 16 – 20.
[28] The non-negligible frequency of positive feedback for very low wages is due to sarcasm. Coders were instructed to code what was literally said rather than trying to infer the true meaning. When an employee sends a message of "Thanks!" following a wage of 0, this is coded as positive feedback even though we doubt the employee was expressing genuine gratitude.

explanation of the high wages when messages are possible. Any such explanation must also show why wages are originally no higher with messages than without, but rise steadily over time. The content of employees' messages can help us understand what might drive these dynamics.

One possible source of learning is that employers don't initially anticipate the strong positive (negative) responses to high (low) wages. As they learn the positive (negative) response that high (low) wages are likely to produce, they adjust wages over time to make a positive response more likely relative to negative responses. This effect could persist even after messages are no longer possible if employers care about the emotions conveyed by the messages rather than the messages per se.[29] If the observed increase in wages is driven by learning about the verbal and/or emotional responses to wages, we would expect to see a positive relationship between lagged receipt of positive feedback following a high wage or negative feedback following a low wage and the current wage.

Alternatively, learning could reflect employers coming to understand that higher wages will yield higher effort. Choosing a wage close to 100 is always the money maximizing action, yet employers in Rounds 1 – 10 of the *Messages Second* sessions (i.e. employers who have never received messages) consistently choose wages that average little more 50. 15 of the 36 employers in this ten round block *never* tried a wage of 90 or greater. In other words, a sizable fraction of employers never tried the payoff maximizing wage. Assuming we don't have a large cohort of deliberately spiteful individuals, it seems likely that these employers didn't realize the profitability of choosing a higher wage. Advice can play a role in helping employers to understand the relationship between wages and effort. If so, we would expect to see a positive relationship between lagged receipt of advice and the current wage.

The regressions contained in Table 3 are designed to separate these explanations for learning. The wage is the dependent variable in all of these regressions. Because employers were restricted to choose a wage between 0 and 100, inclusive, the data is censored. A tobit model is therefore used. The standard errors are corrected for clustering at the employer level.

Model 1 is designed to confirm our observation that wages increase significantly over time when messages are possible, but not when messages are unavailable. The base is Rounds 1 – 10 of the *Messages Second* sessions – the data where employers have no experience with messages. As independent variables, dummies are included for the other three cells of the dataset (*No Messages* * Rounds 11 – 20, *Messages* * Rounds 1 – 10, and *Messages* * Rounds 11 – 20). The independent variable "Rounds in Current Block" measures the round in the current block of ten rounds. In other words, this

---

[29] In other words, employers might care that employees are happy with a high wage or unhappy with a low wage. The messages may help them learn how happy or unhappy the employees are following a specific wage.

variable corresponds to the round for Rounds 1 – 10, resets to a value of 1 in Round 11, and counts up to 10 again through Rounds 11 – 20. The final independent variable is an interaction between "Rounds in Current Block" and a dummy for whether messages are possible (*Messages).*

The parameters of interest in Model 1 are those associated with "Rounds in Current Block" and its interaction with *Messages*. The estimate for "Rounds in Current Block" is small, negative, and nowhere close to statistical significance. There is no evidence that employers move towards higher wages when messages are unavailable. The interaction term is large, positive, and statistically significant at the 1% level. Consistent with our observation from Figure 1, wages are strongly increasing over time when messages are possible.[30]

Model 2 examines the relationship between the employees' messages and increasing wages. Because lagged variables are being used, data from Round 1 is dropped. The lagged wage must be included as an independent variable. There are obvious individual effects in the data and the other lagged variables are all correlated with the lagged wage. Omitting the lagged wage would therefore substantially bias the estimates for the other lagged variables. With the inclusion of the lagged wage as an independent variable, the estimates for the other variables should be interpreted as measuring their effect on ***changes*** in the wage rather than levels. Once again, the base is Rounds 1 – 10 of the *Messages Second* sessions and dummies are included for the other three cells of the dataset (No Messages * Rounds 11 – 20, Messages Possible * Rounds 1 – 10, and Messages Possible * Rounds 11 – 20). "Round in Current Block" is included as an independent variable to capture any changes in wages that reflect pure experience effects. Model 2 also includes the interaction between "Round in Current Block" and "*Messages"* to control for the possibility that pure experience effects differ across the two treatments. Deleting this interaction effect has little impact on the results of Model 2.

The independent variables of interest in Model 2 are the lagged variables reflecting choices of the employee the employer was matched with in the previous round. The parameter estimate for the lagged effort is negative and significant at the 5% level while the estimate for the interaction between lagged wage and lagged effort is positive and significant at the 1% level.[31] These results are consistent with

[30] The dummies for "*Messages* * Rounds 1 – 10" and "*Messages* * Rounds 11 – 20" are positive but not statistically significant. This does *not* contradict our conclusion that allowing messages increases wages, but instead can be attributed to the inclusion of controls for "Round in Current Block" and "*Messages* * Round in Current Block". If these two variables are dropped, the dummies for "*Messages* * Rounds 1 – 10" and "*Messages* * Rounds 11 – 20" are both significant at the 1% level. The effect of allowing messages on wages is not immediately present, but only develops over time.
[31] The variable used in the regression is actually the lagged wage multiplied by the lagged effort and divided by 100. This scaling makes the parameter estimate roughly the same magnitude as other estimates for display purposes.

reinforcement learning.[32] Intuitively, the likelihood that an employer chooses a wage similar to the previous round's choice is increasing in his payoff from the previous round. Since effort and payoffs are directly linked, this implies that the likelihood that an employer chooses a wage similar to the previous round's choice is increasing in the effort from the previous round If an employer chose a low wage in the previous round, he is more likely to choose a low wage in the current round if the previous round's effort was relatively high. Likewise, an employer is more likely to stick with a high wage if it led to a relatively high effort. Neither lagged positive nor lagged negative feedback has a significant effect on the current wage, but lagged receipt of advice has a strong positive effect on current wages which is statistically significant at the 1% level. Model 3 reports an equivalent regression to Model 2 with the dataset limited to observations where messages are possible. Data from Rounds 1 and 11 are dropped to allow for use of lagged variables. The results parallel those of Model 2. In particular, neither positive nor negative lagged feedback has a significant effect on the current wage, but there is a very strong, positive, and statistically significant effect from lagged advice.

We have run alternative versions of Models 2 and 3 that include interactions between the lagged wage and lagged positive feedback, lagged negative feedback, and lagged advice. The interaction terms between lagged wage and lagged positive and negative feedback capture the idea that the effects of positive and negative feedback should depend on the wage. For the modified Model 2, none of the three new interaction terms are statistically significant at even the 10% level (although the interaction between lagged advice and lagged wage barely misses).[33] In the modified version of Model 3 the interaction term between lagged advice and lagged wage edges up to statistical significance at the 10% level. Not surprisingly, advice to choose a higher wage matters slightly more if the employer isn't already choosing a high wage. The other two interaction terms once again fail to approach statistical significance.[34]

The results of Models 2 and 3 reveal the source of increasing wages with messages possible. Given that experiencing positive and negative messages has a negligible effect on wages, there are no grounds to conclude that employers are learning about the verbal or emotional responses of employees to wages. In contrast, advice plays an important role in generating increased wages. The defining characteristic of advice in our experiment is that employees tell employers that a higher wage will lead to lead to higher effort. When messages are not possible, employers who choose low wages have no way of finding out

[32] See Sutton and Barto (1998) for a recent survey of the literature on reinforcement learning.

[33] The parameters for the interaction terms between lagged wage and lagged positive feedback, lagged negative feedback, and lagged advice are .099, .009, and -.364 respectively. The respective standard errors are .179, .243, and .222.

[34] The parameters for the interaction terms between lagged wage and lagged positive feedback, lagged negative feedback, and lagged advice are .018, -.114, and -.420 respectively. The respective standard errors are .197, .269, and .223.

what would have happened if a high wage had been chosen. With messages, the employees can tell them. Having this counterfactual helps employers realize that the positive response of effort to wages makes high wages profitable. Once an employer has learned this, he doesn't forget it and the effect of having messages possible persists even if employees can no longer send messages.

Conclusion 3: Wages are higher when messages are possible, consistent with Hypothesis 3, but the underlying mechanism does not appear to rely on positive or negative feedback. Instead, the primary factor driving increasing wages (and by extension effort and total surplus) when employee messages are possible is advice from employees to employers.

5. **Discussion:** The purpose of our study was to examine whether allowing a second channel of reciprocation beyond effort changes the nature of gift exchange between employers and employees. Allowing employees to send messages at the same time as they choose an effort does indeed change the nature of gift exchange, but not in the predicted manner. The theory predicts that allowing employee messages should affect the effort-wage curve, but no significant change is observed. The theory assumes that any change in wages will be in response to changes in incentives due to shifts in the effort-wage curve plus a desire to receive verbal rewards and avoid verbal punishments. Wages are higher when employee messages are allowed, but this effect is due to employers receiving advice that helps them learn the benefits of high wages. The theory gives us no reason to expect a continuation of high wages once the possibility of employee messages is removed. In reality the effect of allowing employee messages is quite persistent. Once an employer has learned the value of setting a high wage, he doesn't forget it even though he is no longer receiving messages from employees. The overarching conclusion is that employee messages matter in our experiment because they allow employees to advise their employers.

One reason we were interested in studying the effect of employee messages was as a possible explanation for why gift exchange varies across settings. Our experiment supports the possibility that ease of communication can explain some of the variation, but not in the way we anticipated. Allowing messages does not affect gift exchange in the sense that effort is more or less sensitive to wages. What allowing messages affects is the likelihood that employers take advantage of strong gift exchange by setting high wages. The key role of communication in our experiments is not as a second channel of reciprocation, but instead as a channel that allows information to flow from employees to employers. Employees have all sorts of valuable information, including information about their preferences. Communication can give employers access to this information. In settings where it is easy for employees to communicate with their employers and, even better, where giving constructive advice is encouraged, we expect employers to be more likely to take advantage of profitable opportunities for gift exchange. In settings where communication is difficult, gift exchange may play a lesser role, not because gift exchange is unprofitable but rather because it is unconsidered.

The positive role of advice from employees to employers in our experiments is reminiscent of strong evidence from field settings that productivity is increased by Japanese style management systems which, among other things, foster communication between workers and managers (Ichniowski and Shaw, 1999). Our experiments reproduce a similar phenomenon in the controlled environment of the lab. It would be a stretch to claim that our study explains the benefits of a Japanese style management system – we are studying a stylized environment designed to get at a specific set of issues, and exclude many important elements of a Japanese style management system – but we think it reasonable to state that our experiments provide direct evidence of one mechanism by which these management practices are likely to improve productivity.

Taking a broader point of view, any institution that increases the ability of employers to learn about what would have happened if a different wage had been chosen should make it more likely that employers will take advantage of profitable opportunities gift exchange. Possible examples include observing outcomes from other employers or hiring a knowledgeable consultant. Given the benefits of mutually profitable gift exchange, we believe that research identifying what types of institutions make it easiest for employers to learn about employees' responses to different wages will be quite valuable.

Our paper adds to the growing evidence of the importance of advice for learning. The existing result most closely related to ours is reported by Brandts and Cooper (2007). They study a setting where a manager tries to get a group of employees to coordinate in a weak link game by changing their incentives to coordinate and sending messages to the employees. Brandts and Cooper find that successful coordination is more likely if the employees can also send messages to the manager because employees often give the manager useful advice. The incentive structure in Brandts and Cooper is quite different from the BGE and the problems facing their managers differ substantially from those facing our employers. Nonetheless, the underlying conclusion is much the same. Often employees have valuable information about what a manager (employer) should do. If both parties can gain from sharing this information, allowing the employee to advise the manager can be quite beneficial. Other recent experimental papers illustrating the value of advice for learning include Schotter and Sopher (2003, 2006, and 2007), Chaudhuri, Schotter, and Sopher (2009), and Cooper and Kagel (2011).

Our study was designed to emphasize the role of verbal rewards and punishments. This led to our choices to only allow communication from employees to employers and to use a (modified) absolute strangers matching rather than a fixed matching. We have little doubt that communication will have effects beyond those observed here if communication is allowed from employers to employees and/or if a fixed matching is used. If an indefinitely repeated version of the BGE game is played, we conjecture that bilateral communication will assist coordination on a cooperative equilibrium.[35] Use of a fixed matching also raises the possibility that employees will communicate strategically about their likely responses to a high wage. Reputation building and strategic manipulation of information played little role in our experiments by design.

There is surprisingly little evidence that employers in our dataset respond to non-monetary rewards and sanctions from employees despite the fact that these types of messages are common. Wages are not significantly higher when employee messages are introduced nor do positive and negative feedback play important roles in raising wages. This lack of an effect contrasts with the results of Masclet et al. (2003) and Noussair and Tucker (2005) who find strong effects from non-monetary rewards and punishments. We speculate that the difference lies in the nature of payoffs in our experiment. Because of a high efficiency factor and strong gift exchange, employers' profit-maximizing choice of wage is (almost) 100. The interests of employers and employees conflict in theory, but in practice they are closely aligned. For employers, there is no inherent tension between doing the "selfish thing" and doing the "right thing". This eliminates the need for verbal sanctions to "correct" selfish behavior in favor of more socially responsible choices. Further work is clearly needed on when non-pecuniary rewards and punishment will serve as substitutes for monetary rewards and punishments.

[35] For studies showing that communication increases cooperation in repeated games see Isaac and Walker (1988), Ostrom, Walker, and Gardner (1992), Sally (1995), Cason and Mui (2009), and Cooper and Kuhn (2010). This is only an incomplete list of the many studies showing the relationship between communication and cooperation.

References

Akerlof, George. 1982. "Labor Contracts as Partial Gift Exchange" *The Quarterly Journal of Economics,* 97(4): 543-569.

Al-Ubaydli, Omar, Steffen Andersen, Uri Gneezy, and John List. 2008. "For love or money? Comparing the effects of non-pecuniary and pecuniary incentive schemes in the workplace." Working paper. George Mason University.

Andreoni, James, William Harbaugh and Lise Vesterlund. 2003. "The Carrot or the Stick: Rewards, Punishments, and Cooperation." *The American Economic Review*, 93(3): 893-902 –not used.

Ben Greiner: An Online Recruitment System for Economic Experiments. In: Kurt Kremer, Volker Macho (Eds.): Forschung und wissenschaftliches Rechnen 2003. GWDG Bericht 63, Göttingen: Ges. für Wiss. Datenverarbeitung, 79-93, 2004.

Blount, Sally. 1995. "When Social Outcomes Aren't Fair: The Effect of Causal Attributions on Preferences". *Organizational Behavior and Human Decision Processes* 63 (August): 131-44.

Bolton, Gary and Axel Ockenfels. 2000. "ERC: A Theory of Equity, Reciprocity, and Competition." *American Economic Review*, 90(1): 166-193.

Brandts, Jordi and David J. Cooper. 2007. "It's What You Say, Not What You Pay: An Experimental Study of Manager-Employee Relationships in Overcoming Coordination Failure". *Journal of the European Economic Association*, Vol. 5(6), pp. 1223-1268.

Brandts, Jordi and Gary Charness. 2004. "Do Labour Market Conditions Affect Gift Exchange? Some Experimental Evidence." *The Economic Journal*. 114(497): 684-708.

Cason, Tim and Vai-Lam Mui. 2009. "Coordinating Resistance through Communication and Repeated Interaction," Purdue and Monash Universities, working paper, 2009.

Charness, Gary, Guillaume R. Frechette and John H. Kagel. 2004. "How Robust is Laboratory Gift Exchange?" *Experimental Economics*, 7(2):189-205

Charness, Gary. "Attribution and Reciprocity in an Experimental Labor Market." *Journal of Labor Economics*, 2004, vol. 22(3), pp. 665-688.

Chaudhuri, Ananish, Andrew Schotter, and Barry Sopher. 2009. "Talking Ourselves to Efficiency: Coordination in Inter-Generational Minimum Effort Games with Private, Almost Common and Common Knowledge of Advice." *The Economic Journal*, 119(534): 91–122.

Alain Cohn, E. Fehr, and L. Goette. 2009 "Fair Wages and Effort: Evidence from a Field Experiment," University of Zurich, IEW working paper.

Cooper, David J. and John H. Kagel. 2011. "When Are Two Heads Better than One?" Florida State and Ohio State Universities, working paper.

Cooper, David J. and Kühn, Kai-Uwe. 2010. "Communication, Renegotiation, and the Scope for Collusion," Florida State University and University of Michigan, working paper.

Cox, James C., Daniel Friedman, and Steven Gjerstad. 2007. "A Tractable Model of Reciprocity and Fairness." *Games and Economic Behavior*, 59(1): 17-45.

Dufwenberg, Martin, and Georg Kirchsteiger. 2004. "A Theory of Sequential Reciprocity." *Games and Economic Behavior,* 47(2): 268 – 298.

Fehr, Ernst, and Armin Falk. 1999. "Wage rigidity in a competitive incomplete contract market." *Journal of Political Economy,* 107(1): 106–134.

Fehr, Ernst, and Klaus Schmidt. 1999. "A Theory of Fairness, Competition and Cooperation." *Quarterly Journal of Economics*, 114(3): 817-68.

Fehr, Ernst, and Simon Gächter. 2000. "Cooperation and Punishment in Public Goods Experiments." *The American Economic Review*, 90(4): 980-994

Fehr, Ernst, Erich Kirchler, Andreas Weichbold, and Simon Gächter. 1998. "When Social Norms Overpower Competition: Gift Exchange in Experimental Labor Markets." *Journal of Labor Economics*, 16(2): 324-351

Fehr, Ernst, Georg Kirchsteiger, and Arno Riedl. 1993. "Does Fairness Prevent Market Clearing? An Experimental Investigation." *The Quarterly Journal of Economics, 108 (2): 437-459.*

Gneezy, Uri and John List. 2006. "Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments." *Econometrica*, 74(5): 1365–1384.

Hannan, R. Lynn, John H. Kagel and Donald V. Moser. 2002. "Partial Gift Exchange in an Experimental Labor Market: Impact of Subject Population Differences, Productivity Differences, and Effort Requests on Behavior."*Journal of Labor Economics,* 20(4): 923-951.

Ichniowski, Casey and Kathryn Shaw. 1999. "The Effects of Human Resource Management Systems on Economic Performance: An International Comparison of U.S. and Japanese Plants." *Management Science,* 45(5): 704-721.

Isaac, Mark and James Walker. 1988. "Communication and Free-Riding Behavior : The Voluntary Contribution Mechanism." *Economic Inquiry*, 26(4): 585–608.

Kagel, John and Mark Owens. 2010. "Minimum wage restrictions and employee effort in incomplete labor markets: An experimental investigation." Journal of Economic Behavior & Organization, 73(3): 317-326.

Kube, Sebastian, Maréchal, Michel André and Puppe, Clemens. 2010. "The Currency of Reciprocity - Gift-Exchange in the Workplace." University of Zurich Institute for Empirical Research in Economics Working Paper No. 377. Available at SSRN: http://ssrn.com/abstract=1160170

Kube, Sebastian, Maréchal, Michel André, and Puppe, Clemens. 2010. "Do Wage Cuts Damage Work Morale? Evidence from a Natural Field Experiment." Institute for Empirical Research in

Economics,University of Zurich, Working Paper No. 471. Available at SSRN: http://ssrn.com/abstract=1547823]

Masclet, David, Charles Noussair, Steven Tucker, and Marie-Claire Villeval. 2003. "Monetary and Nonmonetary Punishment in the Voluntary Contributions Mechanism." *American Economic Review*, 93(1): 366-379.

Noussair, Charles and Steven Tucker. 2005. "Combining Monetary and Social Sanctions to Promote Cooperation." *Economic Inquiry*, 43(2): 649-660.

Ostrom, Elinor, James Walker, and Roy Gardner.1992. "Covenants With and Without a Sword: Self-Governance is Possible." *The American Political Science Review,* 86(2): 404-417.

Sally, David. 1995. "Conversation and Cooperation in Social Dilemmas: A Meta-Analysis of Experiments from 1958 to 1992," *Rationality and Society*, 7(1): 58 – 92.

Schotter, Andrew and Barry Sopher. 2003. "Social Learning and Coordination Conventions in Intergenerational Games: An Experimental Study." *The Journal of Political Economy,* 111(3): 498-529

Schotter, Andrew and Barry Sopher. 2006. "Trust and trustworthiness in games: An experimental study of intergenerational advice." Experimental Economics, 9(2): 123-145.

Schotter, Andrew and Barry Sopher. 2007. "Advice and behavior in intergenerational ultimatum games: An experimental approach." *Games and Economic Behavior*, 58(2): 365-393

Sutton, Richard S. and Andrew G. Barto. 1998. *Reinforcement Learning: an introduction*. Cambridge: The MIT Press.

Urs Fischbacher (2007): z-Tree: Zurich Toolbox for Ready-made Economic Experiments, *Experimental Economics,* 10(2), 171-178.

Xiao, Erte and Daniel Houser. 2005. "Emotion expression in human punishment behavior," *Proceedings of the National Academy of Sciences of the United States of America*, 102(20), 7398-7401.

Xiao, Erte and Daniel Houser. 2009. "Avoiding the sharp tongue: Anticipated written messages promote fair economic exchange," *Journal of Economic Psychology*, 30(3), 393-404.

Appendix A: Experimental Instructions for *Messages First*

General Information

Thank you for being part of our research. You will be participating in a study of labor markets. If you read these instructions carefully you may earn a significant sum of money paid to you privately at the end of the experiment.

Each of you will be randomly assigned to one of two groups: "Managers" and "Employees." Whether you are a manager or employee will be noted on your computer screen.

Each market period will have four stages:

- Stage 1: Each manager assigns a wage to an employee.
- Stage 2: The employee observes the wage received and may type messages to his or her manager.
- Stage 3: The employee chooses the amount of work to provide.
- Stage 4: Each manager receives the amount of work provided by his or her employee and the market period ends.

Details for computing employee and manager earnings will be provided. The computer screen will display the wage in each period, the amount of work provided, and will transfer all the text from an employee to his or her manager. At the end of the round, the computer will tell you the income that you receive for that round, after which the next round will begin. There will be two 10-round sections of the experiment, with each round played for money. **You will never be matched with the same person twice in the same 10-round section.** Also, you will never know the identity of the person with whom you have been matched. Your total income for the experiment will be the sum of the earnings in each of the ten periods plus a show-up fee.

How the Labor Market Works

1. At the beginning of each period, each manager is matched with one employee. The manager chooses a wage for his or her employee and enters it into the computer. Employees must accept this wage.

2. The computer will display the wage given to each employee on his or her screen.

3. No manager will know the identity of the employee with whom s/he has been paired, and no employee will know the identity of the manager.

4. The employee will have 30 seconds during which s/he is free to type any message into the computer which will be sent to his or her manager. The employee must not write profanity, threatening messages, or anything which identifies the employee!

5. The employee will then choose an amount of work to provide and enter it into the computer.

6. The manager will find out how much work was provided by his or her employee when it appears on his screen. Both managers and employees will have their incomes computed for them by the computer.

How Do Employees Calculate Their Income in Each Period?

1. Employee income in Experimental Dollars is determined by the following formula:

Employee Income = (Wage * 5) + (100 – Amount of Work Provided)

2. Employees determine the amount of work by choosing a number between 0 and 100. The lowest amount of work you can choose is 0, 1 is a slightly higher amount, and so on up to 100, which is the highest amount of work.

3. The higher the amount of work you choose, the better it is for your manager.

4. The higher the amount of work you choose, the higher your work-related costs.

How Do Managers Calculate Their Income in Each Period?

1. Manager income in Experimental Dollars is determined by the following formula:

Manager Income = (Amount of Work Provided * 5) + (100 – Wage)

2. Managers determine the wage by choosing a number between 0 and 100. The lowest wage is 0, 1 is higher, up to 100, which is the highest wage.

3. The higher the wage, the better it is for your employee.

4. The higher the wage you choose, the higher your costs.

Practice Exercises:

Employee Income = (Wage * 5) + (100 – Amount of Work Provided)

Manager Income = (Amount of Work Provided * 5) + (100 – Wage)

1. Assume the manager assigns a wage of 80 experimental dollars, and the employee chooses an amount of work of 20.

What is the employee's income? ______________________experimental dollars

What is the manager's income? ________________________experimental dollars

2. Assume the manager assigns a wage of 30 experimental dollars, and the employee chooses an amount of work of 90.

What is the employee's income? ______________________experimental dollars

What is the manager's income? ________________________experimental dollars

During the experiment your income will be calculated in "Experimental dollars" which will be converted into real dollars at the rate of:

400 experimental dollars = $1

You will also receive a $10 show up fee. There will be 20 total periods. Your total earnings for participating in the experiment will be the sum of your earnings in the 20 market periods plus the show up fee. You will be paid privately at the end of the experiment.

Are there any questions?

Appendix B: Proofs of Claims and Propositions

Proof of Claim 1: We need to show that a responder prefers to accept offers when messages are available that left him indifferent without messages. If $\theta_p = 0$ then all offers will be accepted, so we need only consider $\theta_p > 0$. Without messages, the utility of accepting, $u(\text{accept}) = (1-\theta_p)\ln(P - D + s) + \theta_p \ln(P - D)$, is equal to the utility of rejecting, $u(\text{reject}) = (1-\theta_p)\ln(s) + \theta_p \ln(P)$. Note that this equality is only possible when $D < P$. Because utility is increasing in p as $\theta_p > 0$, responders will always send a negative message when possible. Consider the utility at the new acceptance point, $u(\text{accept with message}) = (1-\theta_p)\ln(P - D + s) + \theta_p \ln(P - D + B)$ and the new rejection point, $u(\text{reject with message}) = (1-\theta_p)\ln(s) + \theta_p \ln(P + B)$. We will henceforth show that $A = u(\text{accept with message}) - u(\text{reject with message}) > 0$.

$$A - u(\text{accept}) + u(\text{reject}) = \theta_p \ln(P - D + B) - \theta_p \ln(P - D) - \theta_p \ln(P + B) + \theta_p \ln(P)$$

$$A = \theta_p \ln\left(\frac{(P - D + B)P}{(P - D)(P + B)}\right) = \theta_p \ln\left(1 + \frac{BD}{(P - D)(P + B)}\right)$$

which proves that A is indeed positive, since $D < P$ and $B, D, P > 0$. Therefore, accepting with a negative message results in a strictly higher utility than rejecting with a negative message, as desired.

Proof of Proposition 2: We need to show that for all $\hat{w}$ and $\hat{\theta}_r$, $e^{**}(\hat{w}, \hat{\theta}_r) \leq e^{*}(\hat{w}, \hat{\theta}_r)$. If $\hat{\theta}_r = 0$, then the employee does not want to give any effort and $e^{**}(\hat{w}, \hat{\theta}_r) = e^{*}(\hat{w}, \hat{\theta}_r) = 0$, as desired. Now consider $\hat{w} \in [0,100]$ and $\hat{\theta}_r > 0$. Let m^* and y^* be the employee's and manager's respective payoffs with no message sent, or

$$m^* = 5\hat{w} - e^{*}(\hat{w}, \hat{\theta}_r) + 100$$

$$y^* = 5e^{*}(\hat{w}, \hat{\theta}_r) - \hat{w} + 100$$

If it is now possible to send a message, then because $\hat{\theta}_r > 0$ it will never be optimal to send a message which is neutral or negative. If it is optimal for an employee not to send a message, then $e^{**} = e^{*}$, satisfying the proposition, because the employees optimal decision cannot change if the choice set has not changed. Therefore we need only consider the case where an employee optimally chooses to send a positive message. The proposition is trivially true when $e^* = 100$, so we only consider $e^* < 100$. Therefore, the MRS at (m^*, y^*) is equal to 5, or perhaps greater than 5 if $e^* = 0$. If, when sending a message, the employee also chooses an effort of $e^{*}(\hat{w}, \hat{\theta}_r)$, then $m = m^* - C$ and $y = y^* + B$. The MRS given this choice of effort is strictly greater than 5 because

$$MRS = \left(\frac{(1-\hat{\theta}_r)}{\hat{\theta}_r}\right)\frac{y^* + B}{m^* - C} > \left(\frac{(1-\hat{\theta}_r)}{\hat{\theta}_r}\right)\frac{y^*}{m^*} \geq 5$$

since $\hat{\theta}_r, B, C > 0$ and $\alpha < 1$. Since the MRS at this point is greater than 5, and the rate at which the employee can exchange his payoff for the manager's is 5, the employee is giving too much to the manager and would like to choose a smaller effort if possible, which implies $e^{**}(\hat{w}, \hat{\theta}) \leq e^{*}(\hat{w}, \hat{\theta})$, as desired.

Proof of Corollary 3: We need to show that for all $\hat{w}$, $e^{**}(\hat{w}, \hat{\theta}_r) < e^*(\hat{w}, \hat{\theta}_r)$ if $\hat{\theta}_r > 0$, $B > 5C$, and $C < e^*(\hat{w}, \hat{\theta}_r) < 100$. From the proof of Proposition 2, we need only show it optimal for the employee to send a message, since $e^*(\hat{w}, \hat{\theta}_r)$ is interior. Because $B > 5C$, and $C < e^*(\hat{w}, \hat{\theta})$, it is possible to obtain both $m > m^*$ and $y > y^*$ when sending a positive message, and since $\hat{\theta}_r > 0$, it is indeed optimal.

The following two lemmata will be useful for Propositions 4 and 5.

Lemma 1: *The employee would optimally choose to send a positive message after receiving any wage above* $\underline{w}$.

Lemma 2: *The employee's optimal effort when messages are not available and his optimal effort when message are available are both continuously increasing functions of wage received, with strict monotonicity for those wages where optimal effort is strictly between 0 and 100.*

Proof of Lemma 1: We need to show that any employee with continuously increasing $\theta_r(w)$ for $w > \underline{w}$ and $u\left(\underline{m}, \underline{y}\right) \leq u\left(\underline{m} - C, \underline{y} + B\right)$ would optimally choose to send a positive message after receiving any wage above $\underline{w}$. Assume not. Then there is a $\hat{w} > \underline{w}$ such that the employee optimally does not send a positive message at $\hat{w}$. If $e^*(\hat{w}, \theta_r(\hat{w})) \geq C$ then we have a contradiction because we know $\theta_r(\hat{w}) > 0$, so the employee would like to increase the manager's payoff and he could do that without reducing his payoff by choosing an effort of $e^*(\hat{w}, \theta_r(\hat{w})) - C$ and sending a positive message. This increases manager payoff because $B > 5C$, which must be the case since a positive message is weakly preferred at the cutoff and if this were true while $B \leq 5C$ then there would be a violation of the convexity of the preferences expressed by the utility function.

Therefore we only need to look at $e^*(\hat{w}, \theta_r(\hat{w})) < C$. Sending a neutral or negative message cannot be optimal, and we have assumed a positive message is not optimal. Therefore, given a wage of $\hat{w}$, it must be that no message and an effort of $e^*(\hat{w}, \theta_r(\hat{w}))$ is preferred to a positive message and zero effort. We will show using Figure A-1 that this contradicts our assumption that $u\left(\underline{m}, \underline{y}\right) \leq u\left(\underline{m} - C, \underline{y} + B\right)$, which will prove the lemma.

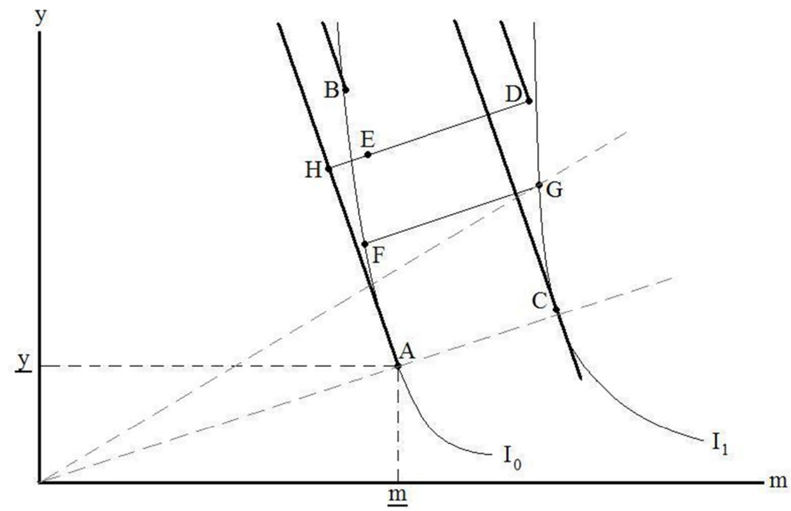


Figure A-1

Point A is the no message cutoff point represented in (m, y) payoff space, and I_0 is the indifference curve tangent to A. B is the same outcome as A with a positive message sent. The bold line extending from A is the employee's choice set with no messages, and the bold line extending from B is the choice set with a positive message. By assumption, B ≽ A as shown. Point C is the outcome for a wage of $\widehat{w}$ and an effort of $e^*(\widehat{w}, \theta_r(\widehat{w}))$ with no message sent, and I_1 is the indifference curve tangent to C. Point D is the outcome for a wage of $\widehat{w}$, effort of 0, and a positive message. The two bold lines are the employee's choice set in each case. By assumption, C ≽ D as shown. Point H is chosen so that $\overline{DH} \perp \overline{AH}$, and point E is the point on $\overline{DH}$ such that $\overline{BE} \parallel \overline{AH}$.

We now demonstrate that the indifference curves drawn in Figure 1 cannot be consistent with our model, because for a given θ_r, the utility function is homothetic, or in other words the MRS depends only the ratio of inputs, in this case, $MRS = \left(\frac{(1-\theta_r)}{\theta_r}\right)\frac{y}{m}$. Thus there is a ray from the origin through any point which contains the set of points with the same MRS. We know that C must be weakly above the ray through A, because as w increases to $\widehat{w}$, θ_r weakly increases, and the ray for which $MRS = 5$, which defines point C, becomes weakly steeper.

Consider the following mapping of points from I_1 onto I_0: an arbitrary point on I_1, say point G in Figure A-1, maps to point F where point F is the point on the line through G which is perpendicular to $\overline{AH}$ which intersects I_0. Because utility is homothetic and θ_r weakly increasing, we know that every point on I_1 above point C (for example G) maps to a point on I_0 (for example F) with a *higher* MRS, because the ray through that point intersects I_0 at a point below the point to which it is mapped (see the dashed line through G). This implies that I_0 is steeper than I_1 for the section of the curves between A (or C) and $\overline{DH}$, which implies that the distance between the indifference curve and the employee's choice set is always greater for I_0 than for I_1. But we have a contradiction because it could not be that C ≽ D unless A ≽ E, but this contradicts the fact that B ≽ A since that would imply E ≻ A given the convex preferences. We have therefore shown that it could not be the case that there exists a wage above $\underline{w}$ for which an employee does not optimally send a positive message, as desired.

Proof of Lemma 2: We need to show that $e^*\big(w, \theta_r(w)\big)$ and $e^{**}\big(w, \theta_r(w)\big)$ are continuously increasing in w, with strict monotonicity when $0 < e < 100$. If the employee is giving full effort for a wage, $\widehat{w}$, then he will also give full effort for wages above $\widehat{w}$, as he will become more willing to substitute his payoff for the manager's and he will continue to send a positive message if possible, by Lemma 1. Therefore, if we show that optimal effort is continuously increasing in wage whenever optimal effort is interior, then the claim must be true for all wages because as wage increases 1.) effort stays at zero until the cutoff 2.) once effort is positive, effort could return to zero only if $e^*\big(w, \theta_r(w)\big)$ or $e^{**}\big(w, \theta_r(w)\big)$ is decreasing or discontinuous over some range of wages for which effort is interior, and 3.) effort stays at 100 once it reaches 100.

We now consider the case of interior effort. Since effort is positive, the employee must have received a wage above $\underline{w}$, and by Lemma 1, the employee sends a positive message when possible. We also know that $MRS = 5$, and $\theta_r > 0$. Therefore, when messages are available

$$MRS = 5 = \left(\frac{(1-\theta_r)}{\theta_r}\right)\frac{y}{m}$$

$$\frac{5\theta_r}{(1-\theta_r)} = \frac{y}{m} = \frac{5e^{**} - w + 100 + b}{5w - e^{**} + 100 - c}$$

Which when solving for e^{**} becomes

$$e^{**} = \left(\frac{24\theta_r + 1}{5}\right)w + \frac{600\theta_r - 5c\theta_r - 100 - b + b\theta_r}{5}$$

which implies that, holding θ_r constant, e^{**} is linearly increasing in w since the coefficient on w is clearly positive and by Lemma 1 we know that $b = B$, and $c = C$ for all wages where effort is interior. This prevents a discontinuity from occurring should the employee switch from sending a positive message to not sending a positive message. If we let $b = c = 0$, then the analysis is equivalent to the case where messages are not possible, and in this case we see there is no change in the linear relationship between optimal effort and wage holding emotional state constant. If we instead hold wage constant and consider an increase in θ_r, it is clear that regardless of message sent effort is weakly increasing because higher θ_r, all else equal, implies a greater concern for the manager's material payoff. In both cases optimal effort is a continuous function of w and θ_r. Therefore, as desired, we have[36]

$$\frac{d}{dw}[e^*(w, \theta_r(w))] = \frac{\partial e^*}{\partial w} + \frac{\partial e^*}{\partial \theta_r} * \frac{d\theta_r}{dw} \geq 0$$

$$\frac{d}{dw}[e^{**}(w, \theta_r(w))] = \frac{\partial e^{**}}{\partial w} + \frac{\partial e^{**}}{\partial \theta_r} * \frac{d\theta_r}{dw} \geq 0$$

[36] In this formula we write $\frac{d\theta_r}{dw}$ without having assumed θ_r is differentiable in w. Since $\theta_r(w)$ is assumed continuous, if it turns out $\theta_r(w)$ is non-differentiable for a given wage we could instead perform the same analysis with arbitrary $\Delta w > 0$ to prove the intended results.

with strict inequality for interior levels of effort as $\frac{\partial e^*}{\partial w} > 0$ and $\frac{\partial e^{**}}{\partial w} > 0$, while $\frac{\partial e^*}{\partial \theta_r}, \frac{\partial e^{**}}{\partial \theta_r}$, and $\frac{d\theta_r}{dw}$ are all non-negative.

Proof of Proposition 4: We need to show that for any employee with continuously increasing $\theta_r(w)$ and $u\left(\underline{m}, \underline{y}\right) \leq u\left(\underline{m} - C, \underline{y} + B\right)$, $\underline{w} < \underline{w}'$ if $\underline{w} < 100$ and $B > 5C$. By Lemma 1 and Lemma 2, we need only show there is some wage above $\underline{w}$ at which the employee exerts no effort. When messages are not possible, effort is positive for wages arbitrarily greater than $\underline{w}$, so $MRS = 5$ at this point. As proved in Proposition 2, the MRS increases when a positive message is sent, which means we can find some $\varepsilon > 0$ such that a positive message is sent and $e^{**}\left(\underline{w} + \varepsilon, \theta_r\left(\underline{w} + \varepsilon\right)\right) = 0$, because $\underline{w} < 100$ and by the continuity of MRS in w and θ_r, the MRS at this point is greater than 5 implying that optimal effort should stay at zero, and therefore $\underline{w} < \underline{w}'$ as desired.

Proof of Proposition 5: We need to show that $\frac{d}{dw}[e^*(\widehat{w}, \theta_r(\widehat{w}))] < \frac{d}{dw}[e^{**}(\widehat{w}, \theta_r(\widehat{w}))]$. By Lemma 1 and Lemma 2, $0 < e^{**} < 100$ can only happen with a positive message and $MRS = 5$. If $e^*(\widehat{w}, \theta_r(\widehat{w})) = 100$, then by Lemma 2, the Proposition holds because as w increases, e^* is constant at 100, while e^{**} is increasing. By Proposition 2, if $e^{**} > 0$ then $e^* > 0$. Therefore we need only consider wages where both effort levels are interior. From the proof of Lemma 2, we know that

$$\frac{\partial}{\partial w}[e^*(\widehat{w}, \theta_r(\widehat{w}))] = \frac{\partial}{\partial w}[e^{**}(\widehat{w}, \theta_r(\widehat{w}))]$$

as the effort is linear in wage holding θ_r constant with the same slope in both cases. Since we want to show that

$$\frac{\partial e^{**}}{\partial w} + \frac{\partial e^{**}}{\partial \theta_r} * \frac{d\theta_r}{dw} > \frac{\partial e^*}{\partial w} + \frac{\partial e^*}{\partial \theta_r} * \frac{d\theta_r}{dw}$$

and we know $\frac{\partial e^*}{\partial w} = \frac{\partial e^{**}}{\partial w}$, and also that $\frac{d\theta_r}{dw}$ does not depend on whether messages are allowed, we need only show that $\frac{\partial e^{**}}{\partial \theta_r} > \frac{\partial e^*}{\partial \theta_r}$.

From the proof of Lemma 2, we have the closed form solution for interior levels of effort, which when written out becomes

$$e^{**} = \frac{1}{5}(24w\theta_r + w + 600\theta_r - 5c\theta_r + b\theta_r - 100 - b)$$

Thus $\frac{\partial e^{**}}{\partial \theta_r} = \frac{1}{5}(24w + 600 - 5c + b)$ while $\frac{\partial e^*}{\partial \theta_r} = \frac{1}{5}(24w + 600)$. The proposition is thus proved because the difference between these two partial derivates, or $\frac{\partial e^{**}}{\partial \theta_r} - \frac{\partial e^*}{\partial \theta_r}$, is $\frac{1}{5}(b - 5c)$, which must be positive because $b = B$, $c = C$, and $B > 5C$ follows from Assumption 1 as proved in Lemma 1.

Table 1: Descriptive Statistics

Wages

	Rounds 1 – 10	t-test	Rounds 11 - 20
Messages First	78.6	4.182 ***	87.2
t-test	3.604 ***		2.217 **
Messages Second	57.0	4.760 ***	73.5

Effort

	Rounds 1 – 10	t-test	Rounds 11 - 20
Messages First	42.6	0.356	44.0
t-test	2.040 **		0.886
Messages Second	28.3	2.307 **	36.3

Employer Payoff

	Rounds 1 - 10	t-test	Rounds 11 - 20
Messages First	234.4	0.117	232.7
t-test	2.706 ***		1.079
Messages Second	184.7	2.145 **	208.0

Employee Payoff

	Rounds 1 - 10	t-test	Rounds 11 - 20
Messages First	450.3	5.617 ***	491.8
t-test	7.843 ***		3.436 ***
Messages Second	356.5	6.390 ***	431.3

Note: Three (***), two (**), and one stars (*) indicate the difference of means is significantly different from zero at the 1%, 5%, and 10% levels.

Table 2: Tobit Regressions on the Effort-wage Curve

	Model 1	Model 2	Model 3
Constant	-89.04*** (15.50)	-69.37*** (14.72)	-70.77*** (13.98)
Rounds 11 – 20	-16.15*** (6.07)	-56.70*** (14.55)	
Messages	-0.19 (6.22)	-12.91 (11.73)	
No Messages * Rounds 11 - 20			-41.20 (27.74)
Messages * Rounds 1 - 10			-5.99 (20.31)
Messages * Rounds 11 - 20			-75.61*** (19.84)
Wage	1.62*** (0.22)	1.32*** (0.22)	1.30*** (0.19)
Wage * Rounds 11 – 20		0.52*** (0.18)	
Wage * *Messages*		0.21 (0.18)	
Wage * Rounds 11 – 20 * *No Messages*			0.42 (0.36)
Wage * Rounds 1 – 10 * *Messages*			0.19 (0.33)
Wage * Rounds 11 – 20 * *Messages*			0.80*** (0.26)
Log-likelihood	-3838.43	-3833.17	-3832.38

Notes: All regressions include 1440 observations from 72 employees. Three (***), two (**), and one (*) stars indicate statistical significance at the 1%, 5%, and 10% levels respectively. Standard errors, reported in parentheses, have been corrected for clustering at the employee level.

Table 3: Tobit Regressions on Wages and Lagged Messages

	Model 1	Model 2	Model 3
Dataset	All Data	All Data Round 1 Excluded	*Messages,* Rounds 1 & 11 Excluded
# Observations	1440	1368	648
Constant	69.36*** (8.96)	-.085 (7.36)	-2.40 (8.75)
Messages * Rounds 1 – 10	16.23 (13.02)	6.14 (6.53)	
Messages * Rounds 11 – 20	5.30 (9.60)	3.70 (5.83)	-4.98 (6.21)
No Messages * Rounds 11 – 20	81.71*** (16.09)	26.27*** (7.76)	
Round in Current Block	-0.76 (1.09)	-0.53 (0.62)	1.12 (0.81)
Messages * Round in Current Block	5.77*** (1.56)	0.89 (0.92)	
Lagged Wage		1.05*** (0.11)	1.09*** (0.12)
Lagged Effort		-0.41** (0.18)	-0.29 (0.26)
Lagged Wage * (Lagged Effort/100)		0.96*** (0.23)	0.82*** (0.29)
Lagged Positive Feedback		-3.89 (5.03)	0.08 (5.41)
Lagged Negative Feedback		6.24 (9.35)	10.68 (9.48)
Lagged Advice		28.13*** (7.09)	30.33*** (7.72)
Log-likelihood	-3775.54	-3047.70	-1443.99

Notes: All regressions include observations from 72 employers. Three (***), two (**), and one (*) stars indicate statistical significance at the 1%, 5%, and 10% levels respectively. Standard errors, reported in parentheses, have been corrected for clustering at the employer level.

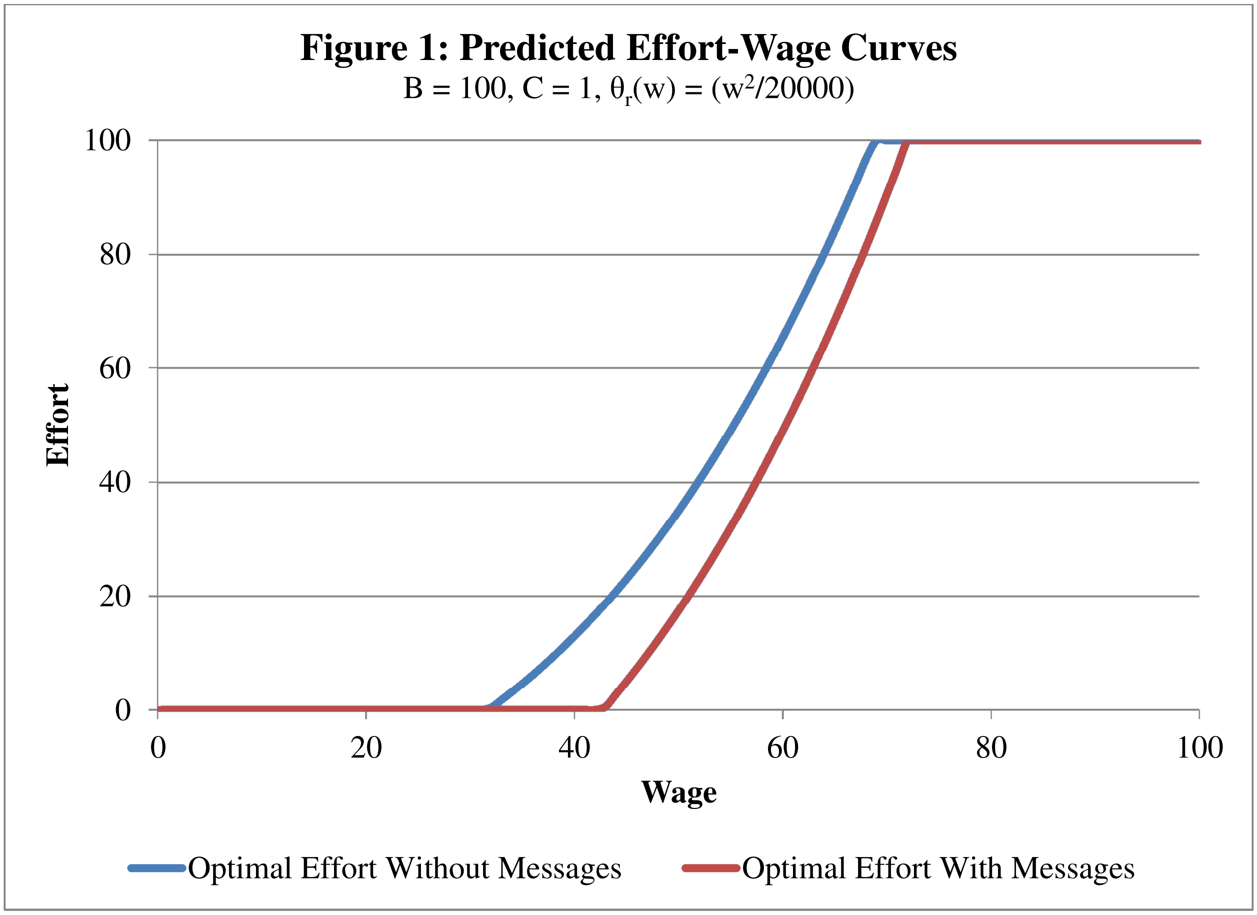
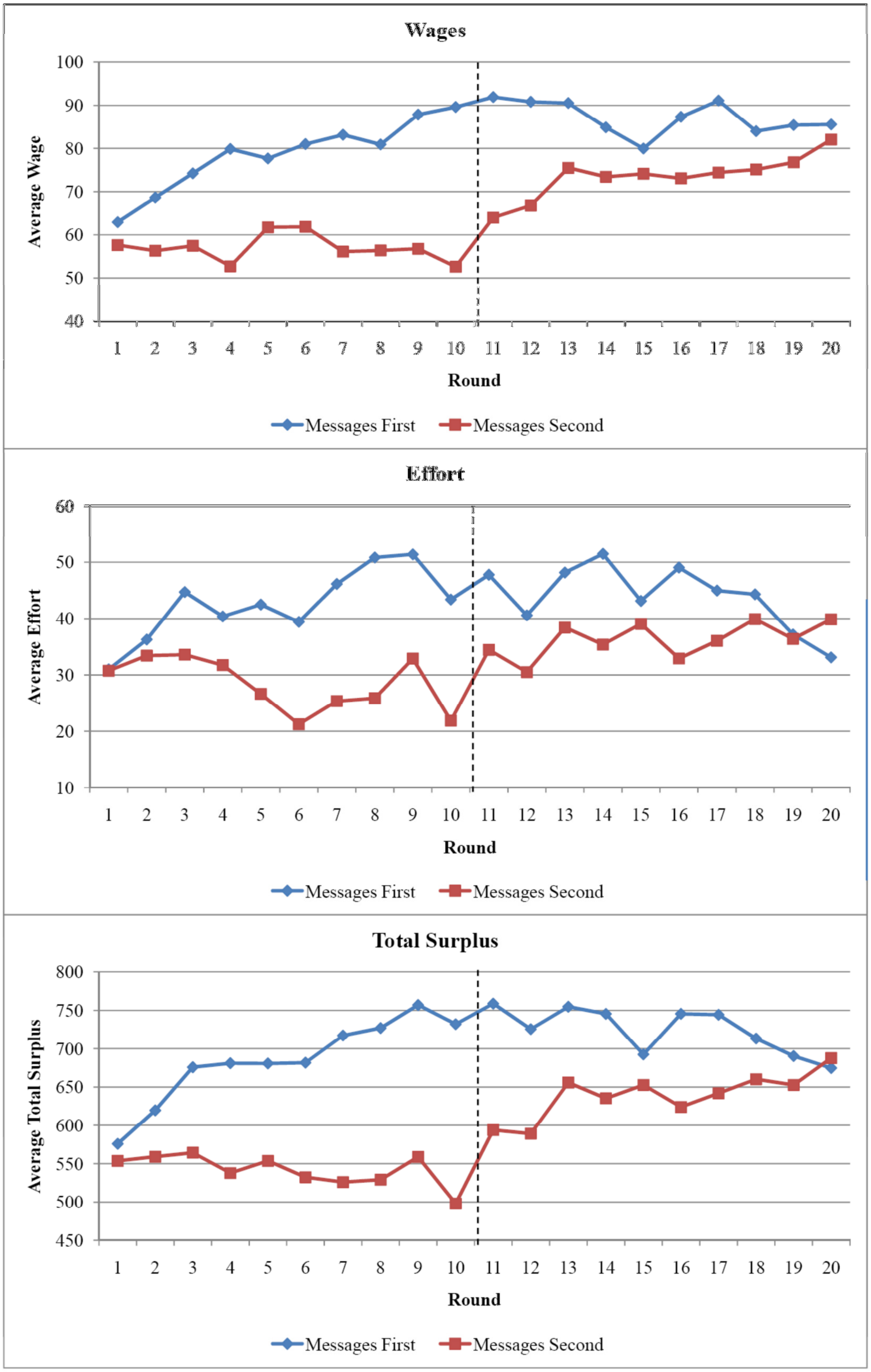


Figure 1: Predicted Effort-Wage Curves
B = 100, C = 1, $\theta_r(w) = (w^2/20000)$
Effort
100
80
60
40
20
0
0
20
40
60
80
100
Wage
Optimal Effort Without Messages
Optimal Effort With Messages

Figure 2: Wage, Effort, and Surplus by Treatment



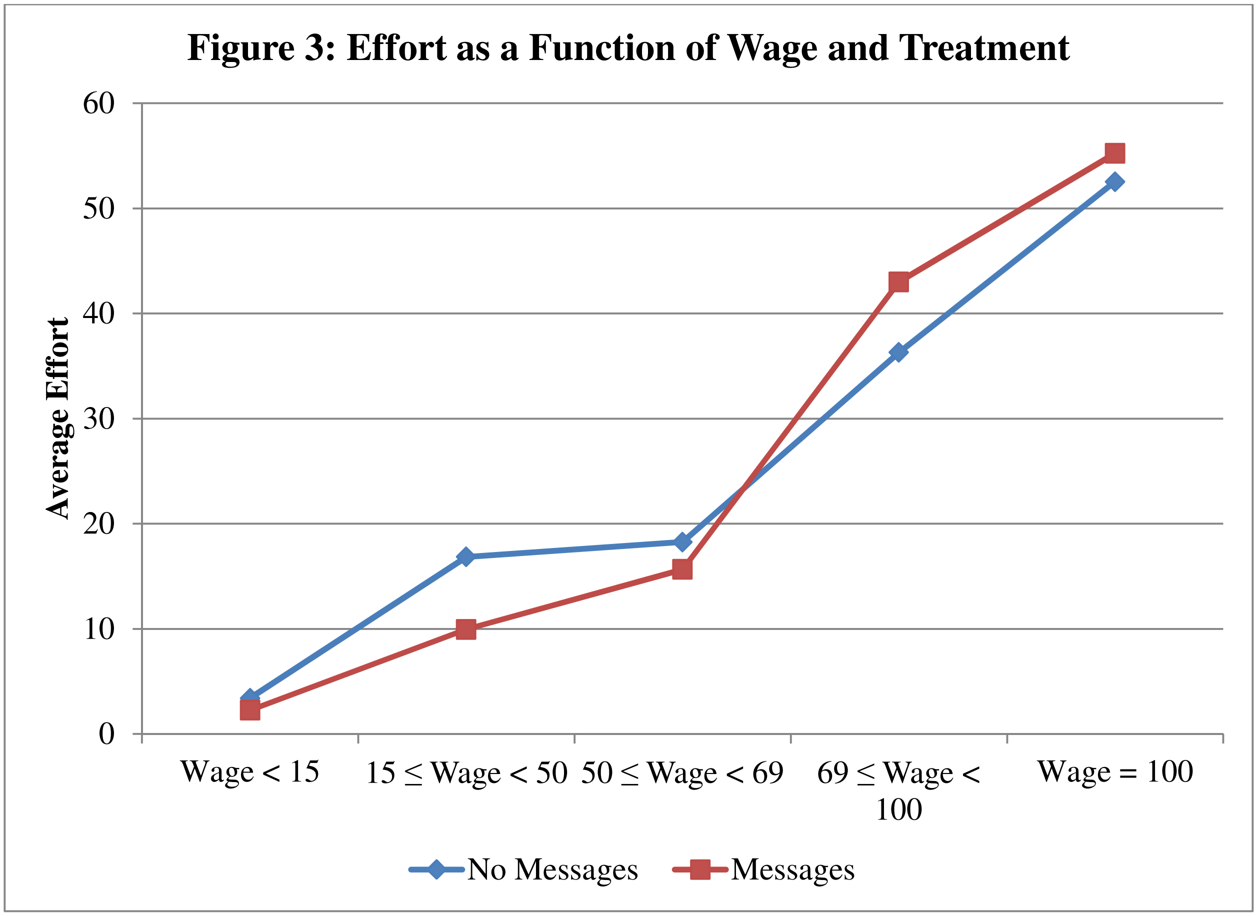


Figure 3: Effort as a Function of Wage and Treatment
Average Effort
60
50
40
30
20
10
0
Wage < 15
15 ≤ Wage < 50
50 ≤ Wage < 69
69 ≤ Wage < 100
Wage = 100
No Messages
Messages

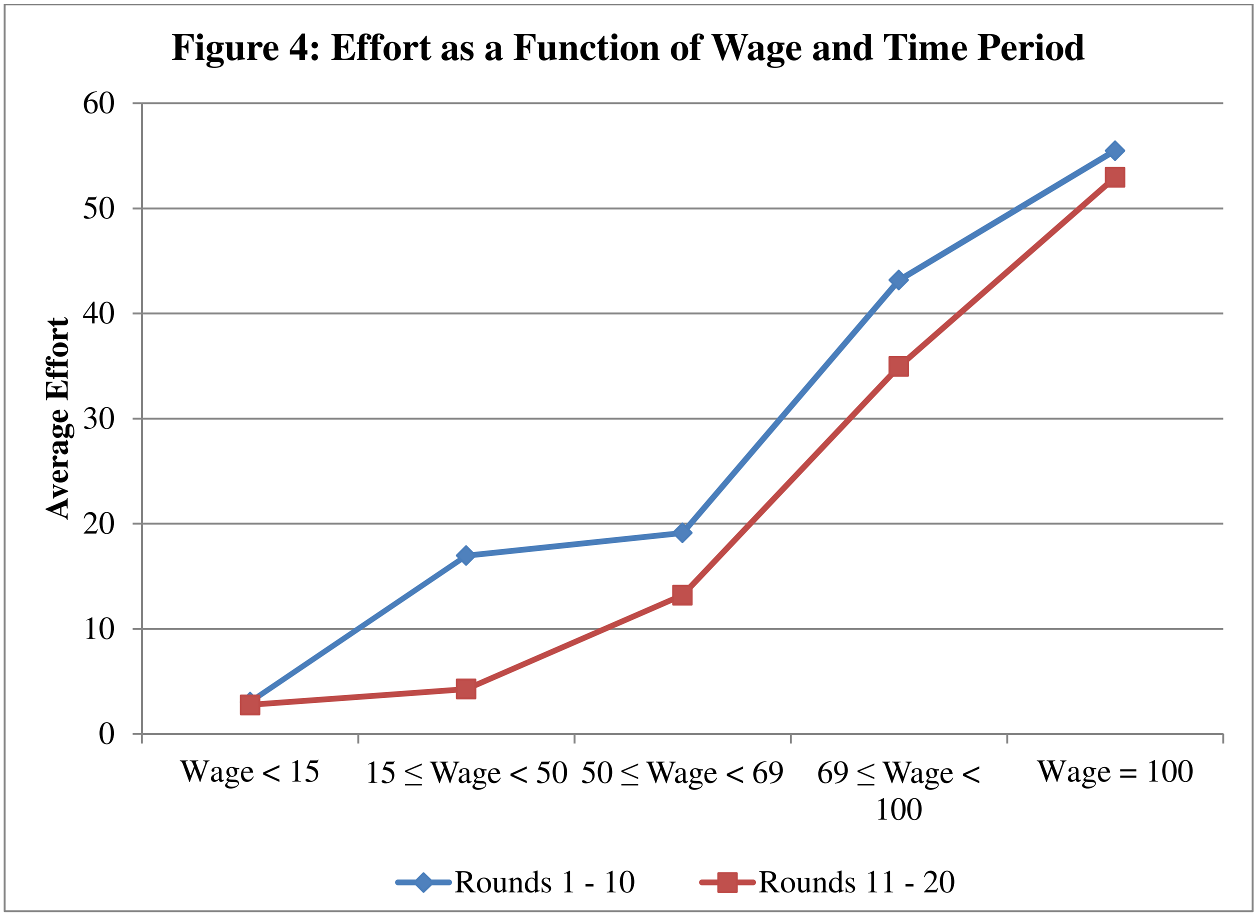
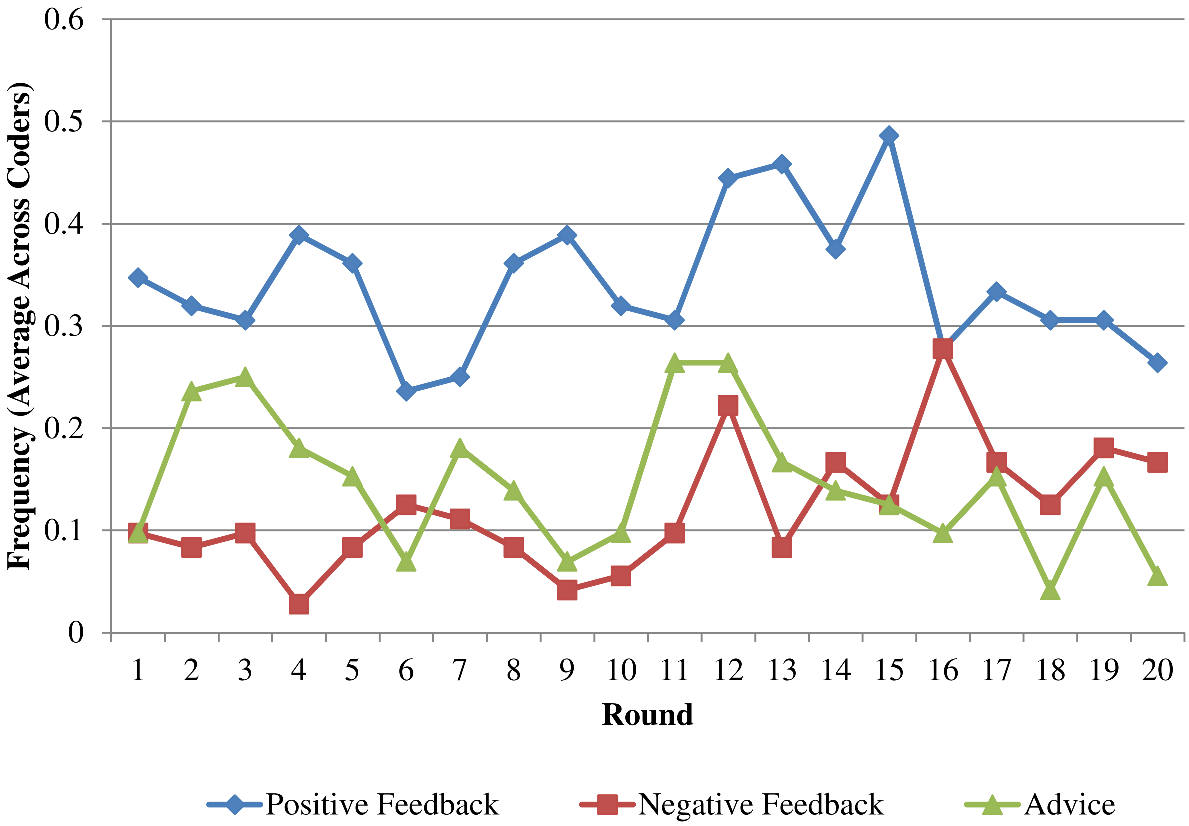


Figure 4: Effort as a Function of Wage and Time Period
Average Effort
0
10
20
30
40
50
60
Wage < 15
15 ≤ Wage < 50
50 ≤ Wage < 69
69 ≤ Wage < 100
Wage = 100
Rounds 1 - 10
Rounds 11 - 20

Figure 5: Frequency of Comments when Messages Are Possible



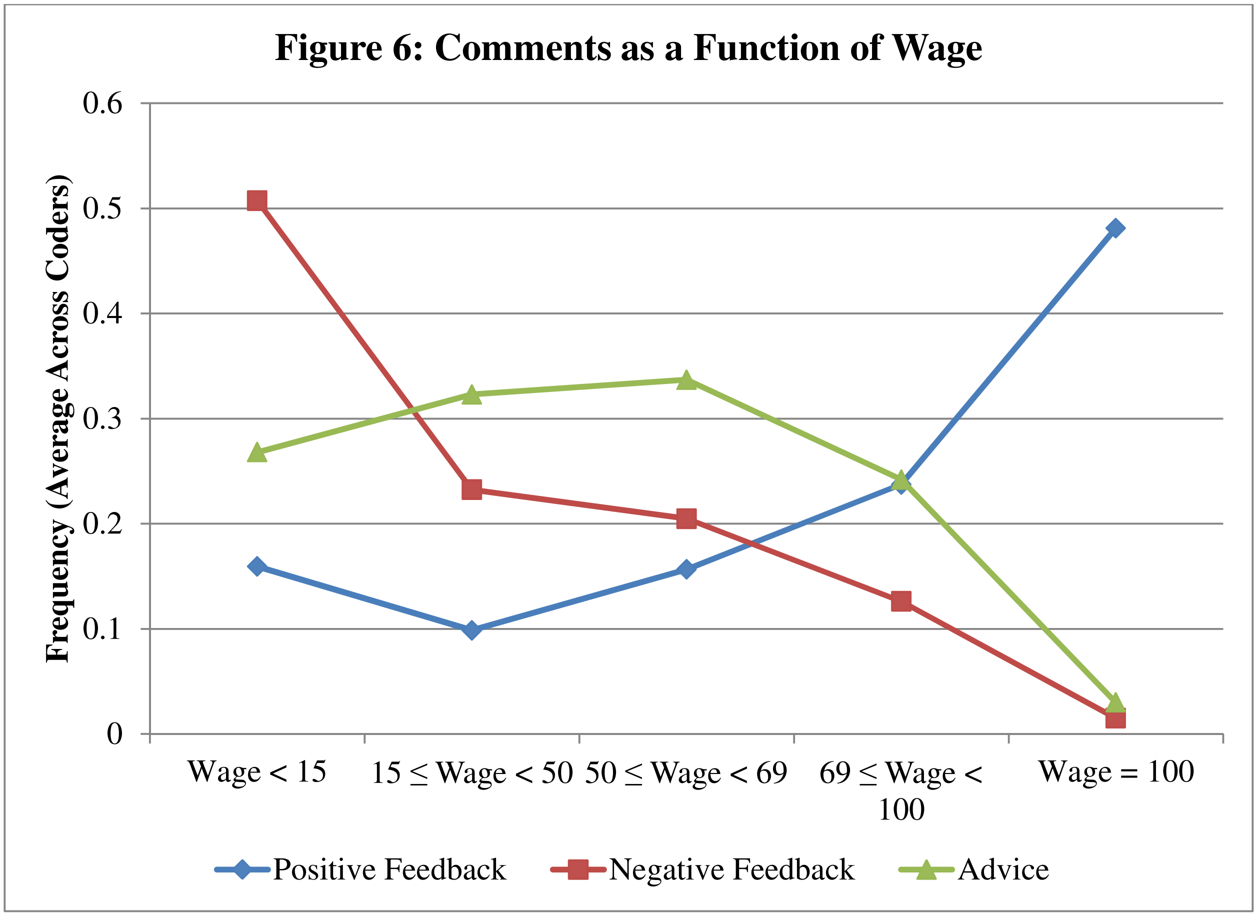


Figure 6: Comments as a Function of Wage
Frequency (Average Across Coders)
0.6
0.5
0.4
0.3
0.2
0.1
0
Wage < 15
15 ≤ Wage < 50
50 ≤ Wage < 69
69 ≤ Wage < 100
Wage = 100
Positive Feedback
Negative Feedback
Advice