

# Performance Pay and the Erosion of Worker Cooperation: Field Experimental Evidence

Stephen Burks  
University of Minnesota, Morris  
svburks@mrs.umn.edu

Jeffrey Carpenter  
Middlebury College  
& IZA  
jpc@middlebury.edu

Lorenz Götte  
University of Zurich  
& IZA  
lorenz@iew.unizh.ch

January 5, 2005

Note: this is an initial rough draft. Comments are most welcome,  
but please to do not cite any specific findings without consulting the authors first.

**Abstract:** We report on the results of a field experiment with bicycle messengers in San Francisco and three Swiss cities, framed at courier workplaces. Messengers work in a setting in which individual output is identifiable enough that firms can choose to condition pay on it, but in which there are still important externalities to messenger behavior that give their on-the-job behavior a social dilemma character. Firms are thus subjected to efficiency losses when messengers fail to cooperate. Second-mover behavior in our sequential strategic form Prisoner's Dilemma allows us to identify *Egoists* (those who defect unconditionally), *Conditional Cooperators*, and *Altruists* (those who cooperate unconditionally). First, we find that messengers, like our student controls, have heterogeneous preferences, but are much more cooperative than the students. Second, contrary to the standard game theoretic assumption, among our subjects beliefs about the cooperation of others strongly affects second mover cooperation. Third, messengers at firms that pay for individual performance are significantly less cooperative. Last, conditional on our identification of exogenous preferences for selecting into particular firm types, we find that the erosion of cooperation under performance pay is predominantly a "treatment effect" as opposed to the result of selection, and that the treatment effect is relatively rapid, more akin to the differential cueing of an existing behavioral norm than the gradual acquisition of a new preference.

**JEL Codes:** C72, C78, C91

**Keywords:** strategic, sequential, Prisoner's Dilemma, social dilemma, bicycle messengers, egoism, conditional cooperation, altruism, social preferences, field experiment, game theory

## 1 Introduction

Economists have long recognized the extent to which the organization of work affects the economic behavior of the workforce (Armen Alchian and Harold Demsetz, 1972). For most of that time the economics profession has relied on the assumption that individual monetary gain and the avoidance of risk are the primary motivations that determine economic behavior. Under this assumption, pay based on work time or on team output in a teamwork setting can dilute the incentives faced by individual workers. A standard recommendation of labor economists in response to this problem is that employers should pay for individual performance, when individual performance is sufficiently measurable, to better align the incentives of individual employees with the goals of the employer (Edward Lazear, 2000). However, even when production is individualized enough that performance pay can be used, the remaining implicit teamwork features of the work process can cause interactions between employees to have a social dilemma character. This creates the potential for a countervailing work process inefficiency—the pursuit of work strategies that are individually beneficial but collectively inferior by selfishly rational agents—and raises the question of whether performance pay might actually increase such behavior, creating a trade-off between the positive and negative effects of this method of compensation.

In addition, recent laboratory experiments have caused economists to begin to rethink the link between monetary incentives and motivation. Now there is substantial evidence that in laboratory subjects there is a potentially complex interaction between responses to individual material incentives and behavior driven by social norms or motivated by social preferences such as reciprocity (Ernst Fehr and Urs Fischbacher, 2002) (Armin Falk and Ernst Fehr, 2002). A fundamental finding of this new literature is that members of all subject pools in experiments in which social preferences can be identified are heterogeneous. Some subjects always behave according to the conventional economic model of rationality, while others are willing to be reciprocal, cooperative, or altruistic, and so on, at a net material cost to themselves. Two basic questions then arise: (1) to what extent does this heterogeneity between selfishly rational and non-selfishly rational agents exist in real world settings of economic interest, and (2) to what extent are the differences in observed behavior driven by the self-selection of people with different preferences into such real-world settings, versus by “treatment effects” of the settings, such as normative cues or learning specific to the setting.

In both the received view, and in the new social preferences view, then, it is relevant to ask whether individual performance pay decreases cooperation in a setting with teamwork characteristics. In this paper we provide empirical evidence from bicycle messenger workers, addressing whether working under performance pay affects the distribution of selfishly rational versus cooperative behavior. Our evidence is from a team production setting in which individual production is sufficiently measurable that some employers have chosen to implement performance pay. Specifically, we measure the degree to which performance pay increases the number of egoistically categorized bicycle messengers in a social dilemma experiment (a strategic form sequential prisoner’s dilemma) in which the social framing is provided by the workplace relationships of the subjects. We also provide a link between our field results and traditional laboratory experimental results using student subjects by comparing messenger behavior to that of a student control group under the same protocol.

Bicycle messengers are especially suitable subjects for this enterprise because their work, delivering message parcels within a short time period, is essentially individualized production. So even though there are many distinct sub-tasks involved in the production of one delivery, the resulting individual performance is measurable enough that firms can choose to condition pay on it. However messenger work also has a significant teamwork component. Some deliveries are better than others because they pay at a higher rate, require less effort, or position the messenger

better for subsequent runs, and messengers can take actions to influence the deliveries they are assigned. Mis-assigned deliveries (e.g., when the closest unoccupied messenger does not make a pickup) have spillover effects on the opportunities available to other messengers and on the total revenue accruing to the firm.

Our study has four major findings. First, we find that messengers do exhibit preference heterogeneity, and they are considerably more cooperative than students in our sequential prisoner's dilemma experiment. As in (Jeffrey P. Carpenter et al., 2005), this suggests that the workplace frame matters.<sup>1</sup> Second, we find that, even though the strategic form of our experiment means that in theory subjects need not concern themselves—in their second-mover choices—with whether others defect, in fact beliefs matter a lot. Controlling for other factors, messengers who believe others will cooperate as a first mover are more likely to cooperate as a second mover. Third, we find that the compensation scheme correlates with cooperative behavior. Messengers at firms that have adopted performance pay are less cooperative both in the game and also according to survey measures of on-the-job and off-the-job behavior, as compared to messengers at firms paying an hourly wage or a share of total revenues. Fourth, and perhaps most interestingly, we find suggestive evidence that the preferences of messengers are endogenous to the employer's choice of compensation scheme. Working under performance pay appears to increase egoism. In fact, conditional on accepting our claim that two variables measure the exogenous selection propensity, our regression analysis suggests that this “treatment” effect is large, robust, and dominates any self-selection by egoists into performance pay firms. The effect is also immediate (within the limits of our survey-based measure of tenure length), which suggests its character could be closer to cueing a behavior-guiding norm than to gradually adopting a new preference.

## 2 The Structure and Incentives of the Courier Industry

Bicycle messengers deliver parcels as small as letters and as big as boxes that require the rider to attach a trailer to his or her bike. Among the workers, such an assignment is known as a “tag.” Many of the tags (on average 37%) are ferrying legal documents around the financial centers of the cities in which we conducted the experiment. However, the pay per tag and the effort required to deliver a parcel is variable. In general, the price charged to a sender depends on the number of “zones” the courier needs to cross to make the delivery. Even for commission-paid couriers the relationship between distance and remuneration is not perfect because legal tags, for example, tend to pay more and, in cities like Zürich and San Francisco, messengers often encounter non-trivial hills within the area served by their firm.

Most couriers work on commission. The average commission-paid messenger in our sample earns 45% of the revenue the tag generates for the firm (40% is a common norm). Among the non-commission riders are those who are paid an hourly wage (the average is \$17.71/hour) and those who are members of coops that share revenues based on the proportion of total work hours contributed. On average, the revenue sharing coops pay out 2.4 times a month and share 59% of the revenue.

In addition to variation in pay schemes, there is some institutional variation in the way that tags are allocated to couriers. Most (94%) of couriers receive assignments directly from a central dispatcher over handheld radios or phones. This procedure is known as *allocated dispatch*.

---

<sup>1</sup> However, we also note that an experiment by **Fehr, Ernst; Kosfeld, Michael and Weibull, Jorgen**. "The Game Prisoners (Really) Play," *Institute for Empirical Research in Economics, University of Zurich, Working Paper*. 2003., shows that a small amount of pre-play communication added to a nearly identical protocol is sufficient to make the level of cooperation displayed by students similar to that of our messengers. This provides some insight about how to capture an analogous framing to that of the workplace in the standard laboratory setting, with easily available subjects.

Despite the primary control held by the dispatcher under this system, messengers can significantly affect their assignments at most firms by how they report their work progress and whether they respond to a particular call. The spatial distribution of pickup and delivery locations means that by reporting a tag done either early or late the messenger can affect the dispatcher's information about whether their location is optimal for taking on a particular new tag. Since radios and phones do not always receive calls inside buildings, a messenger can choose to ignore some calls without penalty. And some firms have their radios set up so that each messenger can hear the assignments made to some or all of the others, which provides information that individual messengers can use either cooperatively or strategically. The remaining few couriers claim tags in the *free call* system, wherein a central dispatcher broadcasts each tag to all of the riders and the rider who responds first makes the pickup.

Under both dispatch systems the individual material incentives at firms that pay commission are those of a common pool resource problem, although the degree of the problem is most severe under free call. Couriers who are so moved can choose to directly "cherry pick" (freecall) the best tags, or manipulate the dispatcher's info to make getting good tags more likely (allocated dispatch), regardless of whether or not they are the best suited to make the delivery. At firms that pay hourly or a time-worked based share of total firm revenues, the incentives are those of a public goods game. Couriers who are so moved can choose to lay back and avoid high effort tags, free-riding (literally) on the efforts of their co-workers. As a result the degree of selfishly rational behavior of its messengers can affect the efficiency of the firm because uncooperative behavior can lead to the misallocation of messenger resources over the geographic area served by the firm.

### 3 Our Field Experiment

We used a strategic form of the sequential prisoner's dilemma game (PD) to measure the cooperative predispositions of the bicycle messengers who took part in our experiment. Each messenger filled out a strategy sheet for their choice as a first mover, and then also for their choices, conditioned on first mover actions, as a second mover. They were informed that a coin toss would decide which of their roles would be activated when we made final payouts.

Playing the PD sequentially provides us with a well-defined typology of player strategies. In the simultaneous version of the PD one can not distinguish, for example, between egoists who simply play the dominant strategy and conditional cooperators who defect, but would have cooperated, provided their partner had done so. But by looking at the second mover strategy chosen in the sequential PD one can identify four possible "types": *Egoists* who defect no matter what, *Altruists* who cooperate regardless of what the first-mover has done, *Conditional Cooperators* who cooperate only if the first-mover cooperates, and what we have termed *Wingnuts* who defect if the first-mover cooperates, but cooperate if the first-mover defects. We interpreted "wingnuts" as participants that did not understand the experiment. Only 3% of our messengers fell into this category, and we chose to exclude these eight observations from our analysis.

Another advantage of the strategic form of the sequential version of the PD is that we can measure the importance of beliefs about the first mover. In theory, the sequential strategic form means that any expectations second movers may have formed about the distribution of cooperative types in the population are irrelevant when implementing a conditional strategy. While this is obvious to a theoretician, it may not have been obvious to our participants and therefore, following Charles Manski (2002) we chose to elicit expectations about the cooperativeness of the pool of players from each of our participants. To give our participants the incentive to think about their estimates, they were paid an additional amount of money for being close to the true distribution of choices. These expectations data allow us to identify the effect of beliefs about others on the choice of strategy. For example, players that ignore their expectations

when choosing the egoistic second mover strategy may well be different than those who choose to be egoists expecting there to be few cooperators in the population.

We chose to maximize the amount of information we gathered from each of our participants by using the strategic form and asking them to make choices in both first and second mover roles both because of the type identification issues discussed above, and also because the high effort and expense of field work with a small and widely dispersed subject pool makes every datum valuable. This raises the question of how comparable our results are to those from simultaneous single-role experiments. Charness and Brandts (2000) find that there is little difference between the “hot”, or live-action, and strategic forms of some standard games used in laboratory experiments. But Burks (2003) showed that in a hot trust game (which is closely related to the sequential PD) with student subjects, playing both roles sharply lowered cooperativeness. So the issue of comparability is not clear. However, we are mostly interested in the difference in cooperation among our participants across variations in their work settings. And we have addressed the question of comparability of results to those from the standard student subject pool directly by running a control with University of Zürich students.

We deliberately put a mild frame on our instructions with messenger subjects because we were interested in our participants’ cooperative predispositions at work, specifically. While the exact wording of our instructions can be viewed in the appendix, the key components of the frame were to label the two roles as “messenger A” and “messenger B”, and to refer to the choices as “cooperate” and “not cooperate”. Our explanation to potential subjects, as contained in the consent document required by the protection of human subjects, and verbal explanations given to those inquiring, was that we were interested specifically in studying bicycle messengers. The remainder of the instructions were neutrally worded. Our goal in this approach was to let the subjects bring with them whatever frame is cued by the facts that they are playing with fellow messengers, that the physical location was associated with their work, and that the study was specifically of messengers.

The details of our protocol are as follows. We conducted the experiment at relatively high stakes: mutual cooperation yielded \$60 for the pair on top of \$15 for filling out our survey. The maximum one could earn was \$64 from not cooperating with a cooperator and accurately estimating the distribution of types in the population; the average payout was \$44 in San Francisco and 59CHF in Switzerland.<sup>2</sup> We first handed out surveys to as many messengers as we could find. In San Francisco we dropped surveys at messenger service offices, and then spent ten noon hours hanging around at “the wall” in the financial district, a small public gathering area where the messengers congregate to eat lunch and wait for delivery calls, and ten afternoons in a nearby café to which the owner permitted us to invite messengers who wished to participate. In Switzerland, we went directly to the breakrooms at the offices of the messenger services, or to public areas immediately outside, to distribute surveys and collect participant decisions. As an incentive, messengers we given \$15 for returning completed surveys. When returning the survey, the messengers were asked to stay for an additional 15 minutes to read and fill out the decision sheets for our experiment.

Because it was impossible to gather all our participants at once, we created a protocol that allowed us to collect one observation at a time. Participants read approximately one page of instructions (see the Appendix) and filled out six control questions. Once a participant had answered the questions correctly (s)he was allowed to proceed. We first asked participants to give their expectations about how many of the other participants would cooperate if the respondent cooperated as the first-mover, and then repeated the question for the case in which the respondent did not cooperate as the first-mover. Participants then chose whether to cooperate or not as the first-mover.

---

<sup>2</sup> All experiments were conducted during the period from May through September, 2003.

The second task of each respondent was to submit a strategy as the second-mover. Participants were first asked to give their expectations about how many first-movers would cooperate. They then submitted a strategy: what they wanted to do if the first-mover cooperated, and what they wanted to do if the first-mover did not cooperate.

As we explained to the participants, to generate final payoffs all the responses we collected were pooled and matched at random, with first versus second-mover roles assigned by the flip of a coin. Payoffs were determined by the intersection of the matched responses and the degree to which the expectations of the subjects were correct. We asked them to indicate whether they preferred to pick up their payoffs (in private) at a later pre-specified date and place, or to have their payoffs mailed.

#### 4 Our Participants

We gathered data from 252 messengers; 139 worked in Zürich, Basel or Bern and the remaining 113 worked in San Francisco.<sup>3</sup> Table 1 provides summary statistics from our participant demographics. Our messengers were rather optimistic about the choices of other first-movers. On average, they expected that 70% of the other players would cooperate in the first-mover role. In fact, these expectations are not optimistic enough because 88% of the messengers cooperated in the first-mover role.

Performance pay is by far the norm among our participants; 82% were paid on commission. Even more common than performance pay is allocated dispatch. Only 6% of our participants were dispatched by free call.<sup>4</sup> The average tenure at the current messenger job was a little more than three years. This figure is higher than most people's prior about turnover in the messenger industry but not surprising considering the high agreement we got to the survey question stating that "working as a messenger is more than just a job." Most messengers (86%) see being a messenger as a lifestyle choice that does not take up too much of their time and pays relatively well. While 39% of messengers work full time (more than 34 hours a week), many others work only a few shifts a week (the average is 27.48 hours) and the average annual earnings of a messenger are \$22,411.

As alluded to at the beginning of this section, our sample is relatively well balanced by location; 55% of the respondents were Swiss. Being a courier is a male-dominated occupation. In our sample 85% of the respondents were men. In terms of other standard demographics, our average participant was 30 years old, high school educated, un-married, and about a third were students.

#### 5 A Summary of Behavior in the Experiment

As an overview of our experimental results, Figure 1 illustrates histograms of the distribution of second-mover types in our experiment by compensation scheme. In the left panel we pool all the messengers who are not paid by commission. This sub-sample is split between 70% who share revenues in different cooperatives and 30% who are paid hourly. We pool the revenue sharing messengers with the hourly paid messengers because they face the same individual material incentive to shirk on the job in contrast to the cherry picking incentive of those paid by commission. As one can see, the two histograms have the same basic shape (the

---

<sup>3</sup> In San Francisco, the anonymous counterparts of each subject in the experiment were drawn from a group that matched the firm size of the subject, either messengers employed at small firms, at medium-sized firms, or at large firms. In Switzerland, where all the firms were medium or large, and two were unique to their home city, we grouped messengers within each firm.

<sup>4</sup> We think these percentages reflect the underlying population of messengers. That is, we did not get the impression that we over-sampled commission couriers or under-sampled free call couriers.

modal participant is a conditional cooperator) but it is also clear that there are more egoists and fewer altruists in performance pay firms.

Because our typology of choices as second-movers results in a categorical or nominal measure of messenger social preferences, to test whether the difference in the distribution of types by firm type is significant, we employ the Pearson  $\chi^2$  test. This test determines whether the difference in proportions exceed those expected by chance. In fact, it does not appear that chance can explain the differences we see in Figure 1 ( $\chi^2=5.27$ ,  $p=0.07$ ).<sup>5</sup> Performance pay firms are associated with more egoism and less altruism.

We also we asked our participants about their behavior at work and off the job. Many of their responses correlate significantly with whether or not they were categorized as an egoist in the experiment. Among the significant correlations were, responses to “I try my best to help out others at work” ( $\rho=-0.13$ ,  $p=0.04$ ); “If the weather is bad, I sometimes take it easy” ( $\rho=0.12$ ,  $p=0.07$ ); “I let messengers who work hard know I appreciate them” ( $\rho=-0.15$ ,  $p=0.02$ ); “How often do you give to panhandlers?” ( $\rho=-0.18$ ,  $p<0.01$ ); “How often do you lend small amounts of money to friends” ( $\rho=-0.11$ ,  $p=0.07$ ). Although these are self-reports, they do suggest that our experimental measure of egoism is a valuable proxy for broader behavior.

## 6 Messengers versus Students

Before we dig deeper into the differences in the behavior of our messengers, we would like to assess the degree to which our pooled messenger distribution of types is similar to what one would see in the traditional lab with student participants. To link our field implementation to the more standard experimental literature we ran a session with students enrolled at the University of Zürich. The only difference between the lab implementation and our field implementation was that we elicited all the student responses at the same time, and the frame was “with other students” as opposed to “with other messengers”. However, the protocol was otherwise the same across the two subject groups.

Figure 2 suggests that there are large differences in the distribution of behavior. There appear to be many more egoists and many fewer altruists in the student population. This difference in proportions is highly significant ( $\chi^2=39.82$ ,  $p<0.01$ ). While it is tempting to conclude that our student control experiment has little external validity with respect to our messengers who face the sort of social dilemma modeled by the experiment on the job, we actually prefer a different interpretation of this difference, one that is bolstered by the experiments reported on in (Ernst Fehr, Michael Kosfeld and Jorgen Weibull, 2003). Fehr et al. also conduct a sequential PD experiment with two treatments. In both treatments they sample from the same pool of University of Zürich students that we did, but in one treatment they allowed the participants to communicate face-to-face with each other before choosing strategies. Our student distribution of types is identical to their baseline implementation but, more importantly, our messenger distribution looks very similar to the distribution they elicit in the communication treatment. This suggests both that the difference between our messenger and student distributions may have to do with how communication cues relevant behavioral norms, and it also suggests it may be possible to capture in the lab a framing analogous to that of the workplace, a setting of great interest to economists, with small adjustments.

## 7 Is Egoism Endogenous to the Compensation Scheme?

Re-focusing on the messenger data alone, we see that the most noticeable difference between the two histograms in Figure 1 is the proportion of egoists working under a particular set of incentives. Although egoism is not close to the mode in either distribution and egoistic types

---

<sup>5</sup> The difference is slightly stronger if we use the nonparametric Fisher’s exact test ( $p=0.05$ ).

account for only 4% (2 of 46) of the participants compensated by revenue sharing or hourly pay, there are four times as many egoists (34 of 206) working on commission. The obvious question is what accounts for the fact that there are so many more egoists working on commission? Are social preferences endogenous with respect to compensation schemes (i.e., does working on commission make people more egoistic) or do egoists disproportionately select into performance pay firms?

Before allowing treatment and selection to go “head-to-head,” we will incrementally build an empirical model of the likelihood of a courier being categorized as an egoist in our experiment. Suppose there is a latent unobservable preference for egoism, expressed in the continuous intensity variable  $e_i^*$ , that we measure in our experiment as an indicator variable,  $e_i$  such that:

$$e_i = \begin{cases} 1, & \text{if } e_i^* > 0 \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

If we assume that the treatment effect of performance pay on egoism and the effect of selecting into performance pay because one is egoistic are additively separable, we can estimate an upper bound on the treatment effect of performance pay by a probit regression of  $e_i$  on an indicator variable,  $c_i$  which takes the value 1 when the participant works on commission and a vector of institutional and demographic controls  $X_i$ . More precisely

$$E[e_i | c_i] = \mathbf{b}c_i + \mathbf{g}X_i + \mathbf{e}_i \quad (2)$$

where  $\hat{b}$  is our estimate of the maximum treatment effect of performance pay, not controlling for any self-selection.

To control for the selection of egoists into performance pay firms, we employ the following two-stage estimation. In the first stage we obtain a probit estimate of the likelihood that our participants end up in a performance pay firm based on a vector of exogenous factors,  $Z_i$

$$\Pr[c_i = 1 | Z_i] = \Phi(\mathbf{t}'Z_i) \quad (3)$$

Using the fitted values from this estimate we calculate the inverse Mills (or hazard) ratio,  $h_i$  for each observation using

$$h_i = \begin{cases} \mathbf{f}(\mathbf{t}'Z_i)/\Phi(\mathbf{t}'Z_i), & c_i = 1 \\ -\mathbf{f}(\mathbf{t}'Z_i)/\Phi(1-\mathbf{t}'Z_i), & c_i = 0 \end{cases} \quad (4)$$

where  $h_i$  gives us the expectation of whether messenger  $i$  works on commission given that the chances are greater than  $\mathbf{t}'Z_i$ .<sup>6</sup>

To obtain a consistent estimate for the treatment effect of performance pay in our second stage regression, controlling for selection into performance pay firms, we add  $h_i$  on the right hand side so that  $E[e_i | c_i]$  becomes the probit regression model

$$E[e_i | c_i] = \mathbf{b}c_i + \mathbf{t}'X_i + \mathbf{r}sh_i + u_i \quad (5)$$

---

<sup>6</sup> The inverse Mills ratio appears as the result of finding the moments of the truncated normal distribution which provides the foundation for most selection models. See Greene (2003: chp.22).

where  $r$  is the correlation and  $s$  is the standard deviation from the bivariate normal distribution of  $e$  and  $c$ .

The results of this method of estimating the treatment effect of performance pay on the likelihood of being an egoist are reported in Table 2.<sup>7</sup> In regressions (1)-(3) we incrementally estimate an upper bound on the treatment effect of performance pay. In regression (1) we see that couriers paid on commission are 10% more likely to be egoists and this marginal effect is highly significant ( $p < 0.01$ ). In regression (1) we limit the set of controls,  $X_i$  but each element of the set appears to have some effect on the odds of being categorized as an egoist. First, we do see that participants condition their choice as a second-mover on their expectations of the cooperativeness of the pool of first-movers. The marginal effect is large: participants who believe that first-movers are all cooperative are 42% less likely to be egoists than those who believe there are no cooperators among the first-movers ( $p < 0.01$ ). However, since there is not a great variation in our subjects' beliefs about the cooperation of others (the standard deviation is .23), the effect in our pool is actually modest: a subject increasing his or her expectation of the cooperativeness of others as first movers by one standard deviation is 9.8% less likely to be an egoist, which is of the same order of magnitude as the treatment effect. Despite the fact that second-movers can condition their cooperation on whether their partner cooperates, we interpret this result to mean that our average participant feels less guilty about being selfish if she believes she is surrounded by selfish people. We also see that messengers dispatched by free call are 18% more likely to be egoists ( $p < 0.01$ ). The free call result is consistent with our priors about free call providing the most intense common pool resource incentives in the messengers' decision problem.<sup>8</sup> Lastly, the small coefficient on our *Tenure* variable indicates that all messengers get slightly more egoistic over time ( $p < 0.01$ ). However, the average messenger with 36.48 months of experience is only 1% more likely to be an egoist than the others.

Given we have established that there is likely to be some treatment effect of performance pay on the likelihood of being an egoist, in regression (2) we ask whether messengers in performance pay firms are "treated" gradually over time, as if they were adopting a new preference, or whether their behavioral change is more abrupt. To answer this question we add the interaction of working on commission and one's tenure on the current job to  $X_i$  to parse out any differential effect of tenure on those who work on commission. The addition of the interaction increases the coefficient on the baseline effect of tenure from 0.0003 to 0.002 indicating the average messenger is now 7% more likely to be egoistic than one just starting out ( $p < 0.01$ ), but contrary to our priors, messengers who work on commission do not become more (or less) egoistic over time. Although the coefficient on the interaction is highly significant ( $p < 0.01$ ), the differential effect exactly cancels the baseline effect; it therefore appears that performance pay has an immediate effect on one's social preference for cooperation, but that there is a small erosion of cooperation over time at non-performance pay firms.

In regression (3) we test the robustness of our results by adding more demographic controls to  $X_i$ . The controls soak up some of the variation attributable to beliefs and substantially increase the marginal effect of the free call dispatch system, but they do not change the treatment effect of performance pay by much. Messengers paid on commission are still 12% more likely to be egoistic. As for the demographics, working more hours has an effect similar to tenure on the job. The average courier who works 27.48 hours a week is approximately 3% more likely to be an egoist than a messenger who works very little ( $p < 0.1$ ), a standard deviation shift upward in

---

<sup>7</sup> We also cluster standard errors by firm to account for heteroskedasticity.

<sup>8</sup> If more of the couriers worked under free call, we would also consider looking at the treatment effect of free call on courier social preferences, controlling for selection into free call firms. However, our first stage regression can not get enough "traction" with only 6% of the population using free call.

household income (\$11,310) increases the likelihood of being an egoist by a little more than 2% ( $p < 0.1$ ), and being Swiss increases the likelihood by an appreciable 7% ( $p < 0.1$ ).

Regressions (4) and (5) represent the two stages of our procedure to control for selection into performance pay firms. Regression (4) reports the marginal effects from our first stage probit regression of working on commission on two identifying factors. In our survey we asked each respondent whether (s)he thought that working on commission was good or bad and whether (s)he thought sharing revenues was good or bad. One's *net preference for commission* is the difference between the (Likert scale) responses to these two questions. Histograms of the net preference by compensation scheme appear in Figure 3. As one can see, this measure strongly identifies not working on commission; there are very few people working for an hourly wage or sharing revenues that think that working for a commission would be better than their current situation. At the same time, the distribution of net preferences has much wider support in commission pay firms. Some couriers would prefer revenue sharing, but many more are happy with performance pay.

Figure 4 illustrates our second identifier. In another survey question we asked for responses to the statement, "When another messenger at my firm doesn't play by the rules, I let him or her know I don't like it" as a measure of the degree to which messengers enforce social norms on the job. Figure 4 suggests that *norm enforcement* is much stronger and less variable at non-performance pay firms. Together, these two identifiers account for 23% of the variation in the likelihood of working on commission and are each highly significant in the stage one probit. If one is willing to grant these attitudes the status of measures of exogenous determinants of the motivation to self-select into one type of firm or the other, they provide the basis for our two-step selection correction model.

Regression (5) reports our estimate of the treatment effect of performance pay on courier social preferences for cooperation, controlling for selection into performance pay firms. As expected (given we assumed the effects of treatment and selection to be additively separable), the coefficient on commission pay shrinks, when we add the inverse Mills ratio. However, two factors suggest selection is not a large problem for our analysis. First, the drop in the coefficient on performance pay is small (from 12% to 11%). Second, the coefficient on the inverse Mills ratio is neither large nor significant ( $p = 0.83$ ). Note, this does not mean that there is no need to control for selection. The real test of whether the addition of  $h_i$  to (4) is warranted is whether the residuals from regressions (4) and (5) are correlated. Indeed the correlation is significant ( $p = 0.08$ ) suggesting that the estimated treatment effect in (4) is upwardly biased.

## 8 Conclusion

We conducted a sequential prisoner's dilemma field experiment with workers (bicycle messengers), some of whom are compensated by performance pay, but who work under social dilemma on-the-job incentives. In addition, we ran a separate control experiment with students to link our results to the laboratory literature on social preferences.

First, we find that messengers do exhibit preference heterogeneity, and they are considerably more cooperative than students in our sequential prisoner's dilemma experiment. As in Jeffrey Carpenter, et al. (2005), this suggests that the workplace frame matters, although comparison of our results with those of Ernst Fehr, et al. (2003) suggests that small changes in the framing of experiments with students in the lab can cause the lab results to mimic those of our messengers. Second, we find that, even though the strategic form of our experiment means that in theory subjects need not concern themselves—in their second-mover choices—with whether others defect, in fact beliefs matter a lot. Controlling for other factors, messengers who believe others will cooperate as a first mover are more likely to cooperate as a second mover. Third, we find that the compensation scheme correlates with cooperative behavior. Messengers at firms that have adopted performance pay are less cooperative both in the game and also according to survey

measures of on-the-job and off-the-job behavior, as compared to messengers at firms paying an hourly wage or a share of total revenues. Fourth, and perhaps most interestingly, we find suggestive evidence that the preferences of messengers are endogenous to the employer's choice of compensation scheme. Working under performance pay appears to increase egoism. In fact, conditional on accepting our claim that two variables measure the exogenous selection propensity, our regression analysis suggests that this "treatment" effect is large, robust, and dominates any self-selection by egoists into performance pay firms. The effect is also immediate (within the limits of our survey-based measure of tenure length), which suggests its character could be closer to cueing a behavior-guiding norm than to gradually adopting a new preference.

9 Appendix – Experimental Instructions

You, and another messenger will now participate in a decision-making problem. The other messenger works at «treat» *Neither you nor the other messenger will ever know each other’s identity.* Please read the instructions carefully. The choices you and the other messenger will make *determine how much cash you each will get* in this decision-making problem. The monetary stakes are high so you should consider your choice carefully. In this decision-making problem, we refer to the two parties involved as “messenger A” and “messenger B”. The choices that the two messengers can make are very similar:

- **Messenger A** has two options: Cooperate or Not Cooperate.
- **Messenger B** has the same two options, but chooses after messenger A. So, first, if messenger A cooperates, Messenger B indicates whether he or she will respond by cooperating or not cooperating. Second, if messenger A doesn’t cooperate, messenger B indicates whether he or she will respond by cooperating or not cooperating.

Your choices will lead to different payoffs for you and for the messenger you are paired with (on top of what you earned so far). The choices, and the payoffs they result in, are summarized in the following table:

**Messenger A’s options**

	Messenger A cooperates		or	Messenger A does not cooperate		
<b>Messenger B’s options</b>	Messenger B responds by cooperating	or		Messenger B responds by cooperating	or	Messenger B responds by not cooperating
	Result ↓			Result ↓		Result ↓
	Messenger A gets \$30			Messenger A gets \$45		Messenger A gets \$15
	Messenger B gets \$30			Messenger B gets \$10		Messenger B gets \$15

**What will you do?**

Each messenger will choose what to do in both the messenger A role and the messenger B role, because you could end up in either role. We will also ask you to estimate what the other messengers will do.

At a later date (see the address sheet), after we gather choices from a number of messengers at your firm, we will match you with another messenger and flip a coin to determine whether you will be messenger A or messenger B (the other messenger will assume the other role).

Once you are matched, and assigned a role as an A or a B, the choices that were made by you and the messenger you are matched with will determine how much each of you will earn. We will pay you your earnings. *Neither you nor the other messenger will ever know each other’s identity.*

**Does everyone understand?**

Before we proceed, please take a minute to answer the following three questions. Everyone will be asked the same questions to ensure that all the participants understand the decision exercise. Use the payoff table on the first page to answer these questions.

Question 1: Suppose messenger A decides to cooperate.

(i) How much would messenger B earn if he does not cooperate?  
*Messenger B would earn \_\_\_\_\_ dollars.*

(ii) How much would messenger B earn if he cooperates?  
*Messenger B would earn \_\_\_\_\_ dollars.*

Question 2: Suppose messenger A decides not to cooperate.

(i) How much would messenger B earn if he does not cooperate?  
*Messenger B would earn \_\_\_\_\_ dollars.*

(ii) How much would messenger B earn if he cooperates?  
*Messenger B would earn \_\_\_\_\_ dollars.*

Question 3: Suppose messenger A chooses to cooperate, and messenger B responds by cooperating.

(i) How much does messenger A earn?  
*Messenger A earns \_\_\_\_\_ dollars.*

(ii) How much does messenger B earn?  
*Messenger B earns \_\_\_\_\_ dollars.*

**Decision Sheet ONE (for your Messenger A choice).**

If the coin flip gives you the messenger A role, your choice from this decision sheet will be used. To decide what you want to do consult the table of choices and earnings on first page of the instructions. You have two decisions to make.

**First:** indicate what you think the group of other messengers, one of which you will be paired with, will do as messenger B. You will be paid an additional two dollars if your two estimates are within plus or minus five percent of the actual percentage of what the other messengers choose.

- *If I cooperate as messenger A, \_\_\_\_\_ percent of the other messengers in my group will choose cooperate as messenger B.*
- *If I do not cooperate as messenger A, \_\_\_\_\_ percent of the other messengers in my group will choose cooperate as messenger B.*

**Second: Your Decision as Messenger A:** Please indicate whether you chose “cooperate” or “not cooperate” if you are assigned to be messenger A.

My choice if I am messenger A

*(please choose only one)*

<input type="radio"/> <i>Cooperate</i>	<input type="radio"/> <i>Not cooperate</i>
--	--

**Please double-check your choices on this sheet.**

**Please notify the experimenters if you are done with this decision sheet.**

**You will then receive the second decision sheet.**

**Decision Sheet TWO (for your Messenger B choice).**

If the coin flip gives you the messenger B role, your choice from this decision sheet will be used. To decide what you want to do consult the table of choices and earnings on first page of the instructions. You have two decisions to make.

**First:** indicate what you think the group of other messengers, once of which you will be paired with, will do as messenger A. You will be paid an additional two dollars if your estimate is within plus or minus five percent of the actual percentage of what the other messengers choose.

- *As messenger A, \_\_\_\_\_ percent of the other messengers in my group will choose to cooperate.*

**Second: Your Decision as Messenger B:** Please indicate what you will do if you are assigned to be messenger B.

My choice if I am messenger B

<i>If messenger A cooperates, I will</i>		<i>If messenger A does not cooperate, I will</i>	
<input type="radio"/> <i>Cooperate</i>	<input type="radio"/> <i>Not Cooperate</i>	<input type="radio"/> <i>Cooperate</i>	<input type="radio"/> <i>Not Cooperate</i>

**Please double-check your choices on this sheet.**

**Please notify the experimenters if you are done with the decision sheet.**

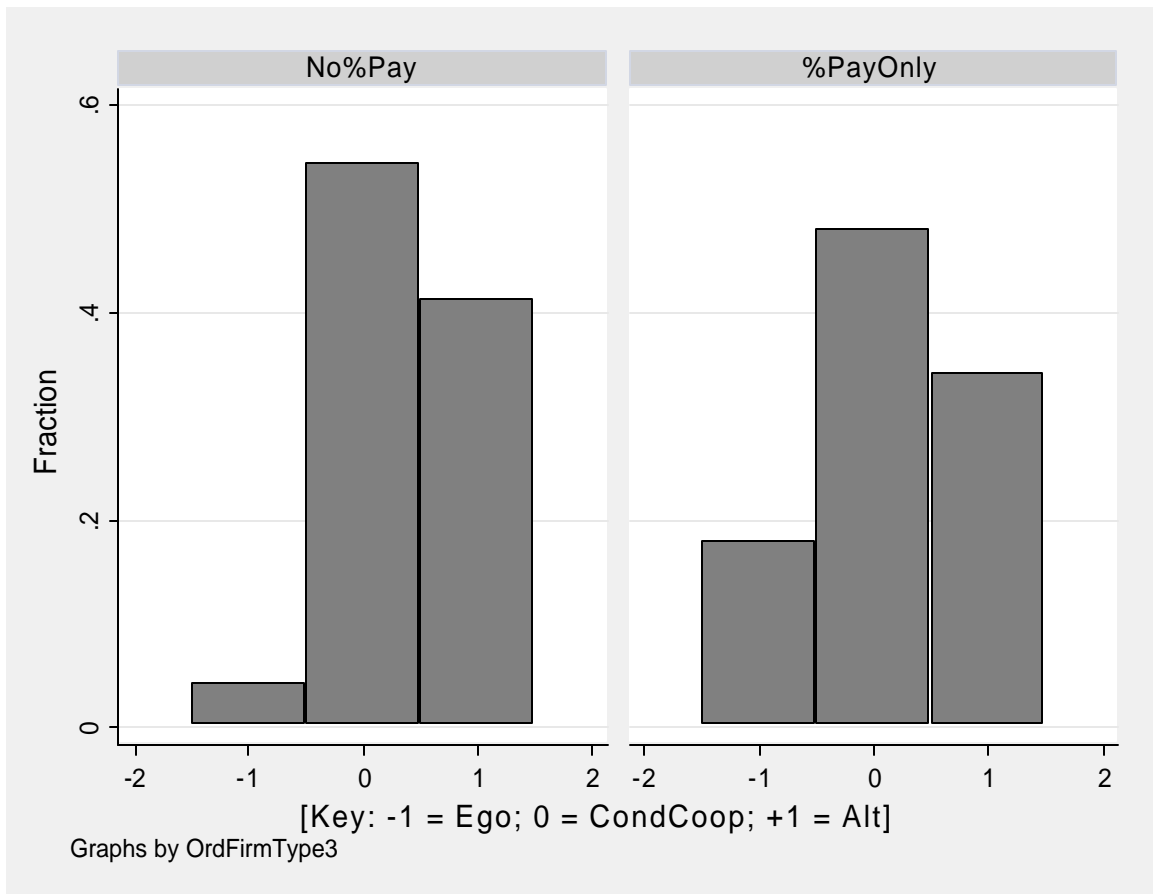
10 Tables and Figures

Table 1: Participant (Messenger) Demographic s					
	Obs.	Mean	Std. Dev.	Min	Max
Beliefs: Fraction(A Cooperate)	252	0.70	0.23	0	1
Paid on Commission	252	0.82	0.39	0	1
Dispatched by Free Call	249	0.06	0.24	0	1
Tenure: months at current job	252	36.48	34.94	0	162
Hours worked as Messenger	250	27.48	14.39	1	80
Household Income (in dollars)	251	22411.35	11310.28	5600	45000
Swiss	252	0.55	0.50	0	1
Male	252	0.85	0.35	0	1
Age	252	29.60	6.47	19	52.5
Married	252	0.15	0.35	0	1
Student	252	0.34	0.47	0	1
Years of Schooling	247	13.85	20.4	9	18

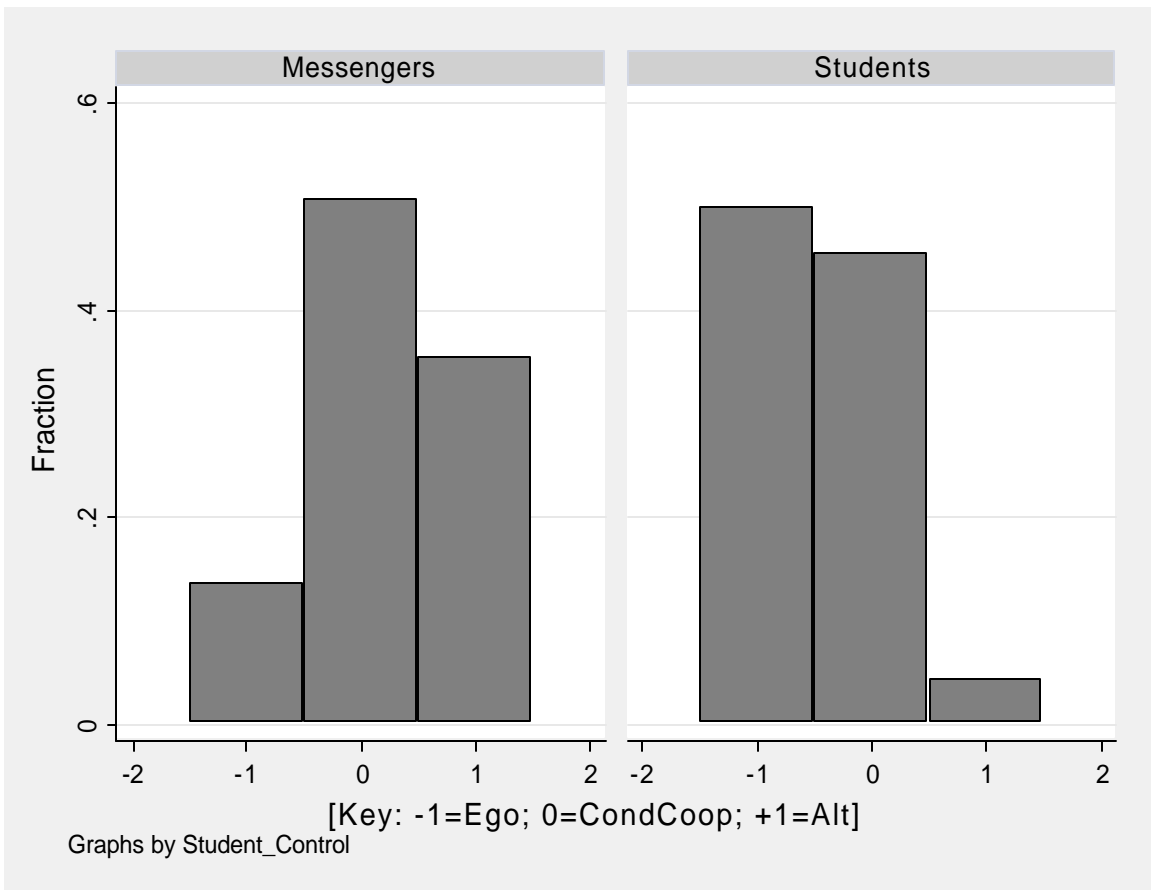
Table 2: Performance Pay and the Likelihood of being an Egoist

	(1)	(2)	(3)	(4)	(5)
Beliefs: Fraction(A Cooperate)	-0.42 (0.08)***	-0.40 (0.08)***	-0.36 (0.10)***		-0.36 (0.10)***
Paid on Commission	0.10 (0.01)***	0.13 (0.02)***	0.12 (0.02)***		0.11 (0.04)***
Dispatched by Free Call	0.18 (0.04)***	0.18 (0.03)***	0.28 (0.04)***		0.29 (0.06)***
Tenure: months at current job	0.0003 (0.0001)**	0.002 (0.0006)***	0.002 (0.0005)***		0.002 (0.0007)***
Tenure × Commission		-0.002 (0.0007)***	-0.002 (0.0006)***		-0.002 (0.0008)***
Hours worked as Messenger			0.001 (0.0009)*		0.002 (0.001)*
Household Income (in thousand dollars)			0.002 (0.000)*		0.002 (0.000)*
Swiss			0.07 (0.03)*		0.08 (0.04)**
Male			-0.02 (0.04)		-0.02 (0.04)
Age			-0.002 (0.003)		-0.001 (0.003)
Married			-0.006 (0.04)		-0.01 (0.04)
Student			-0.02 (0.02)		-0.03 (0.02)
Years of Schooling			0.006 (0.007)		0.004 (0.007)
Inverse Mills ( $h_i$ )					0.01 (0.05)
Net Preference for Commission				0.07 (0.01)***	
Norm Enforcement				-0.06 (0.02)***	
Obs.	249	249	241	249	239
Pseudo R <sup>2</sup>	0.27	0.28	0.31	0.23	0.32

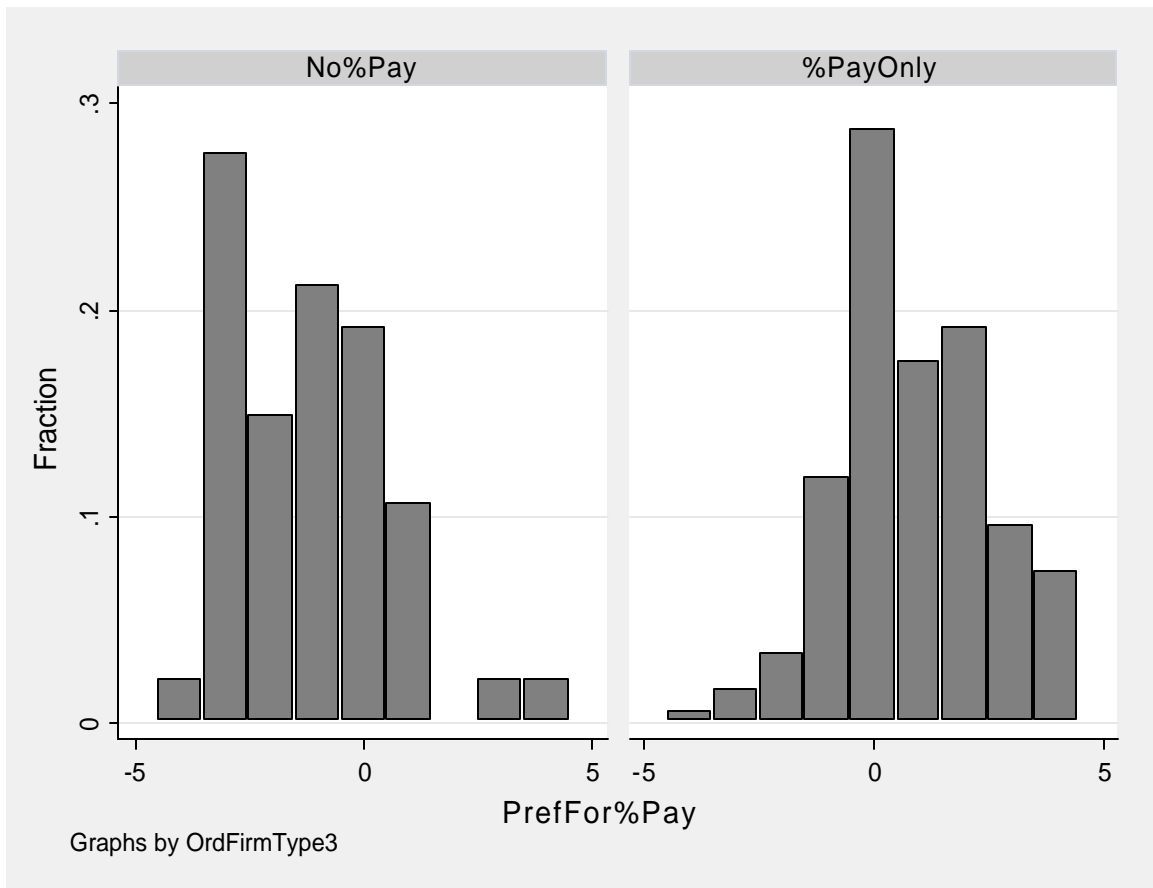
Notes: The dependent variable in regressions 1,2,3,5 is 1 if the participant was categorized as an egoist and 0 otherwise. The dependent variable is regression 4 is 1 if the participant works for commission and 0 otherwise. In each case we report marginal effects after probit regressions. (standard errors). \* p<0.1; \*\* p<0.05; \*\*\*p<0.01.



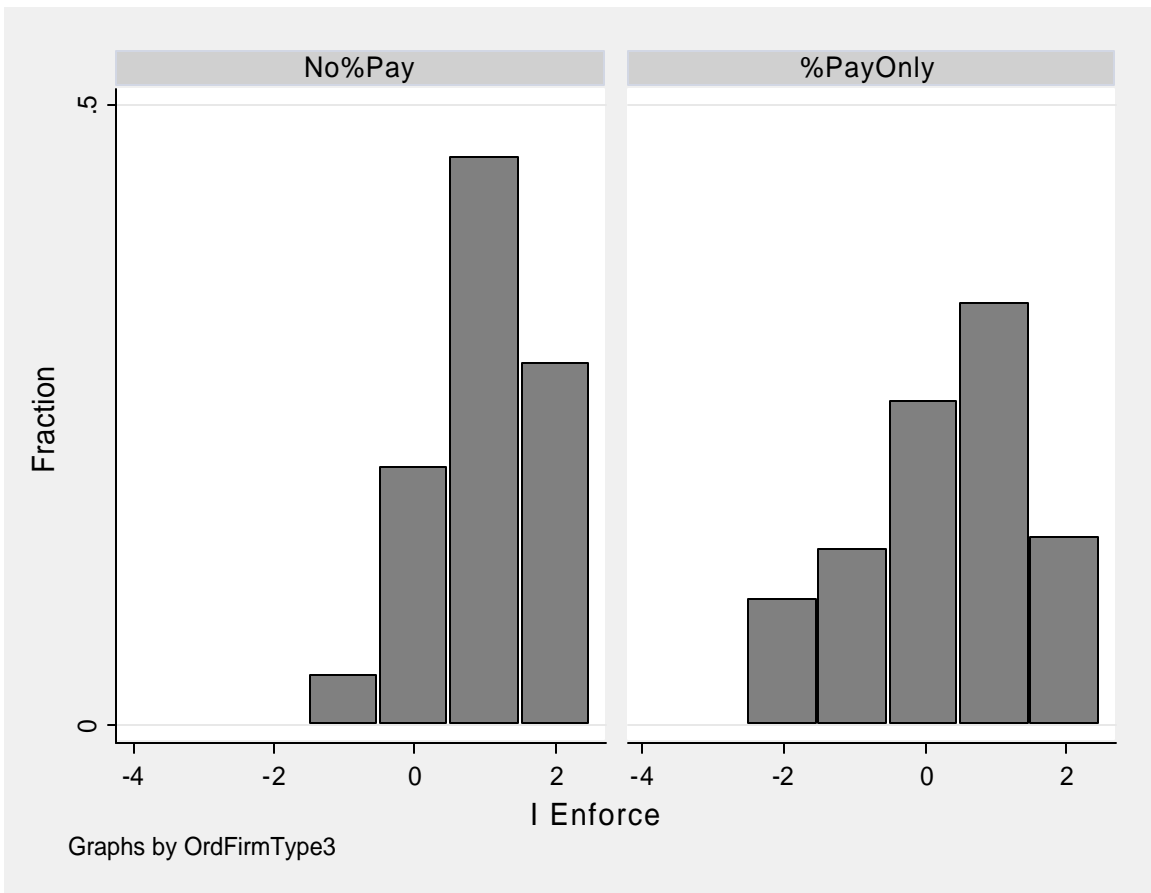
**Figure 1:** The distribution of types in the sequential prisoner’s dilemma by messenger compensation scheme (Note: No%Pay combines hourly paid and revenue sharing messengers. %PayOnly are messengers paid by commission).



**Figure 2:** The distribution of types by field or lab implementation.



**Figure 3:** The distribution of the Net Preference for Commission Pay by compensation scheme (Note: the Net Preference for Commission Pay is the surveyed ordinal preference for working on commission minus ordinal preference for revenue sharing).



**Figure 4:** The distribution of surveyed Norm Enforcement by compensation scheme.

## 11 References

- Alchian, Armen and Demsetz, Harold.** "Production, Information Costs, and Economic Organization." *American Economic Review*, 1972, 62(December), pp. 777-95.
- Burks, Stephen V.; Carpenter, Jeffrey P. and Verhoogen, Eric.** "Playing Both Roles in the Trust Game." *Journal of Economic Behavior and Organization*, 2003, 51(2), pp. 195-216.
- Carpenter, Jeffrey P.; Burks, Stephen V. and Verhoogen, Eric.** "Comparing Students to Workers: The Effects of Social Framing on Behavior in Distribution Games," J. Carpenter, G. Harrison and J. List, *Field Experiments in Economics*. Greenwich, Conn. and London: JAI/Elsevier, 2005,
- Charness, Gary and Brandts, Jordi.** "Hot Vs. Cold: Sequential Responses and Preference Stability in Experimental Games." *Experimental Economics*, 2000, 2(3), pp. 227-38.
- Falk, Armin and Fehr, Ernst.** "Psychological Foundations of Incentives." *European Economic Review*, 2002, 46, pp. 687-724.
- Fehr, Ernst and Fischbacher, Urs.** "Why Social Preferences Matter - the Impact of Non-Selfish Motives on Competition, Cooperation and Incentives." *Economic Journal*, 2002, 112(March), pp. 1-33.
- Fehr, Ernst; Kosfeld, Michael and Weibull, Jorgen.** "The Game Prisoners (Really) Play," *Institute for Empirical Research in Economics, University of Zurich, Working Paper*. 2003.
- Lazear, Edward.** "Performance Pay and Productivity." *American Economic Review*, 2000, 90(5), pp. 1346-61.
- Manski, Charles.** "Identification of Decision Rules in Experiments on Simple Games of Proposal and Response." *European Economic Review*, 2002, 46(4-5), pp. 880-91.