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

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

Jeffrey Carpenter
Middlebury College
& IZA
jpc@middlebury.edu

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

November 4, 2005

Abstract: We report the results of a field experiment with bicycle messengers in Switzerland and the United States. Messenger work is identifiable enough that firms can choose to condition pay on it, but externalities to messenger behavior still exist and give their on-the-job interactions the character of a social dilemma. Firms therefore suffer efficiency losses when messengers fail to cooperate. Second-mover behavior in our sequential Prisoner’s Dilemma allows us to characterize the cooperativeness of our participants. We find that messengers, like our student controls, have heterogeneous social preferences, but are much more cooperative than students. Among messengers, we find that employees at firms that pay for performance are significantly less cooperative than those who are paid hourly or are members of cooperatives. To examine whether the difference is the result of treatment or selection we exploit the fact that firm type is location-specific in Switzerland and that entering messengers must work in performance pay firms in the U.S. This allows us to exogenously identify selection. We find that the erosion of cooperation under performance pay is predominantly due to treatment, and that the treatment effect is relatively rapid, more akin to the differential cueing of a behavioral norm than the gradual acquisition of a new preference.

JEL Codes: C72, C78, C93, D23, J33, J54, Z13
Keywords: sequential, Prisoner’s Dilemma, social dilemma, beliefs, bicycle messengers, social preferences, egoism, conditional cooperation, altruism, field experiment
1 Introduction

It has been quite a while since economists realized the extent to which the organization of work affects the monetary incentives of the workforce Alchian and Demsetz (1972) and for a long time the economics profession has been content in the assumption that monetary incentives fully determined motivation and behavior. However, recent laboratory experiments have caused economists to rethink the link between monetary incentives and motivation. Not only is there substantial evidence showing that subjects often discount material incentives and act according to social norms or preferences like reciprocity Fehr and Fischbacher (2002), they sometimes ignore material incentives altogether and act based on intrinsic motivation Falk and Fehr (2002).

Teamwork in particular characterizes a work process in which the material incentives faced by individual employees can be confounded not only because individual and group monetary incentives draw individuals towards different acts, but because social preferences (preferences over monetary outcomes formed by a preference of one’s own well-being, the well-being of others, beliefs, and intentions) can render material incentives ineffectual. Furthermore, there is potential for a dynamic effect of social preferences if they are endogenous with respect to the organization of work Bowles (1998).

A standard recommendation of labor economists in response to the team production problem is that employers should pay for individual performance, when individual performance is sufficiently measurable, to better align the incentives of individual employees with those of the employer (e.g., Lazear, 2000). However, there are a number of possible problems with this prescription. First, even when production is individualized enough that performance pay can be used, the remaining teamwork features of the work process can cause interactions between employees to have a social dilemma character. Second, such a prescription does not account for the response of socially motivated employees to the implementation of performance pay. Economists have sound predictions for the response of egoists to performance pay, but how will reciprocally, or otherwise socially-motivated agents respond to performance pay? On top of these important issues, and the point of departure for the current study, performance pay itself may partially determine worker social preferences in ways that, at a minimum, have so far been un-modeled, and may be deleterious to production.

We provide empirical evidence addressing whether pay schemes affect the distribution of social preferences in a team production setting in which individual
production is highly measurable and 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 where the social framing is provided by their workplace relationships. To complete the experiment we also provide a link between our field population and traditional laboratory experiments by comparing messenger behavior to a student control group.

Bicycle messengers are especially suitable subjects for this project 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. But even the work of messengers paid by performance has a significant teamwork component because some deliveries are better than others (i.e., pay at a higher rate or require less effort) and messengers can take actions to influence the deliveries they are assigned. That is, miss-assigned deliveries (e.g., when the closest messenger does not make the delivery) create externalities for other messengers and for the total revenue accruing to the firm.

Our research has uncovered three major findings. First, we find that messengers are considerably more cooperative than students in our sequential prisoner’s dilemma experiment. Like Carpenter et al. (2005), this suggests that the workplace frame matters. However, we also note that an experiment by Fehr et al. (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 into how to create an analogous frame to ours in the standard laboratory setting, with easily available subjects. Second, 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 a hourly wage or a share of total revenues. Third, and perhaps most importantly, we find that the social preferences of our couriers are endogenous to the employer’s choice of compensation scheme. Working under performance pay increases egoism. In fact, our regression analysis suggests that this “treatment” effect is large, robust, and dominates any self-selection by egoists into performance pay firms.

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 the 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 total 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. 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 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 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 egoistic 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 typography 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 identity 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. Thankfully, only 3% of our messengers fell into this category, and we chose to exclude these eight observations from our analysis.

In theory, using the sequential form of the PD 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 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 provide us with another dependent variable that might be explained by differences in the employment contract. Not only might the distribution of social preferences be endogenous to the use of performance pay, these differences might correspond to differences in what our messengers expect from each other. For example, it would not be surprising if more egoism evolves under piece rates and the messengers were aware of this in terms of their expectations of how other messengers in their firm will react to social dilemma incentives.

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. Brandts and Charness (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 et al. (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 extent of the frame was to label the two roles as “messenger A” and “messenger B”, and to refer to the choices as “cooperate” and “not cooperate.” The remainder of the instructions were neutrally worded. Our goal in this approach was to let the subjects bring with them whatever frame was cued by the fact that they are playing with fellow messengers, that the physical location was associated with their work, and that the study was 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.\footnote{All experiments were conducted during the period from May through September, 2003.} 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.

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. 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. 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

\footnote{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.}

\footnote{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.}
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 on commission. As one can see, the two histograms have the same basic shape (the 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 typography 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 ($\chi^2=5.27$, $p=0.07$).\(^4\) Performance pay firms are associated with more egoism and less altruism.

In Figure 2 we find the corresponding differences in beliefs that suggest that multiple equilibria will evolve in the distribution of social preferences according to how firms organize compensation. Not only are couriers less cooperative in performance pay firms, they expect that their co-workers to be less cooperative. As an illustration of this, while 25% of the cooperative and hourly workers expect that everyone in their subject pool will cooperate as a first-mover, less than 10% of the performance pay workers hold the same belief. Overall, the distribution of beliefs about the fraction of first-movers who will cooperate is shifted to the right for non-performance pay messengers – the Kolmogorov-Smirnov test is highly significant ($ks=0.24$, $p=0.03$) – indicating their more optimistic expectations.

We also 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 on-the-job.

---

\(^4\) The difference is slightly stronger if we use the nonparametric Fisher’s exact test ($p=0.05$).
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 students. 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 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.” The protocol was otherwise the same across the two subject groups.

Figure 3 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 ($\chi^2=39.82$, $p<0.01$). As within our messenger population, the differences we find in student behavior extend to their beliefs. In Figure 4 we see that the student distribution of beliefs about how many first-movers will cooperate is close to uniform, except that none of our students expected that all the other students would cooperate. It comes as no surprise that these two distributions are highly significantly different (ks=0.41, $p<0.01$) and, again, the pessimistic beliefs of the students correspond precisely to the amount of cooperation they will experience in the experiment.

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 Fehr et al. (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 protocol adjustments.
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 clearly not the mode in either distribution and egoistic types 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. A more subtle difference also exists in the number of altruists. However, relative to the number of conditional cooperators (our basis for comparison), the differences between performance pay firms and the others are smaller. There are 76% as many altruists as conditional cooperators in non-performance pay firms and 71% in performance pay firms. 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?

Two things allow us to identify selection for a sub-sample of our messengers. For the 139 messengers working in Switzerland, differential selection into performance pay is not possible because the employment contract only varies across cities, it does not vary within a city. Hence, all 68 messengers from Zürich and all 29 and 20 from Basel and Luzern, respectively are paid on commission while the 22 working in Bern are members of a cooperative. One might object that couriers could commute to different cities to work under the scheme of their choice, but these distances are non-trivial, and we know for a fact that no one does this. Hence, if a worker decides that he wants to be a courier in Switzerland, he has no choice as to how he will be compensated, and therefore, the compensation scheme is not endogenous for our Swiss messengers.

We can actually still do a little better. In San Francisco, the industry is structured so that new entrants must gain experience in the relatively large performance pay firms before they can decide to join an hourly pay firm or a coop. If a worker decides she wants to be a courier in San Francisco, she will have to start her career as a performance pay employee. This fact gives us another 18 observations for which compensation is not a choice variable. These are workers who have less than 18 months
of experience in the industry. Altogether, 63% of our (non-wingnut) sample is comprised of messengers whose behavior could not be confounded by differential selection into one compensation scheme or another.

The tenacious critic, however, might still object. Even though there is no differential selection of messengers into employment contract for this sub-sample, behavioral types may still differentially select into a career as a messenger and if, for example altruists are more likely to become messengers in Bern and egoists are more likely to become messengers in Zürich, then there is still an issue. Our survey data does not suggest that this is a problem. We asked the messengers to respond (on a 5-point Likert scale from -2 to 2) to the statement, “Working as a bike messenger is more than just a job.” The mean response was 1.33 indicating that most people are very committed to being a bike messenger. Furthermore, the responses to this question do not vary significantly by location ($\chi^2=18.93$, p=0.27) or by compensation scheme ($\chi^2=2.13$, p=0.71) which suggests that our participants become messengers because of lifestyle reasons that have nothing to do with social preference or performance pay.

To assess the magnitude of the treatment effect of performance pay on social preference, we regress the behavioral type of a messenger on the compensation scheme he works under. To be both comprehensive and as agnostic as possible about the differences between types we use the multinomial logit estimator with conditional cooperators as the comparison group and report robust standard errors to account for the possibility of heteroskedasticity.

We build our empirical model gradually by adding controls in four steps. However, there are eight estimates in Table 2 because for each step we compare the estimate using the full sample to the estimate based on only our restricted sub-sample for whom selection is not an issue. These comparisons allow the reader to assess the possible impact of selection at each stage of our analysis. In column (1) we see that, with no controls or conditioning variables, messengers in the full sample are 13% (p<0.01) more likely to be egoists (than conditional cooperators) in performance pay firms. Performance pay messengers are also 4% less likely to be altruists but this result is not significant. The point estimate of the treatment effect of performance pay on

---

5 There are only two exceptions to this rule that make our identification of the employment contract only 99% perfect. That is, there were two new entrants in San Francisco not working under performance pay that had to be recoded as performance pay according to our identification procedure.
egoism falls to 12% (p<0.05) when we shift attention to the restricted sample in column (1r) but the difference is very small and not significant.

As we add controls we see that our estimate of the treatment effect of performance pay on egoism is very robust. While the coefficient hovers between 12 and 15 percent, the standard errors suggest that these fluctuations are not important (i.e., the estimates are within the confidence interval of the other) and therefore that selection is not likely to bias our results much. Equally important, the estimate is always statistically significant.

In columns (2) and (2r) we ask whether the treatment of messengers happens gradually or whether differences in the employment contract cue existing behavioral norms. Because neither tenure on the current job nor the differential affect of tenure in performance pay firms are remotely significant, it appears that social preferences in this case change quickly.⁶

In columns (3) and (3r) we add the eight demographic controls listed in Table 1. For egoists, none of these factors are robust and most are never significant. For egoists, however, one demographic factor is robust. Older messengers are much more likely to be altruistic. A standard deviation increase in age is associated with a 7% increase in the probability of being an altruist in the full sample and a 15% increase in the restricted sub-sample. The stylized fact, that older people appear more prosocial has been replicated in a number of different social dilemma field settings (e.g., List, 2004 or Carpenter and Seki, 2005a).

In sum, we find little evidence of selection as an explanation for the differences in the distribution of social preferences shown in Figure 1. Our estimates of the treatment effect of performance pay on egoism from the sub-sample for whom selection is not possible are very similar to our estimates from the full sample. Performance pay appears to make messengers between 12 and 15% more likely to behave egoistically towards their co-workers.

8 What about beliefs?

⁶ This is somewhat different than Carpenter and Seki (2005a) who find that the social preferences of Japanese fishermen change more gradually in response to differences in the compensation scheme.
Beliefs also appear to depend on the firm’s choice of compensation scheme. As depicted in Figure 2, performance pay workers are (rightfully) less optimistic about how many first movers will cooperate. In this sense, our beliefs data, along with the fact that behavioral treatment appears to happen quickly, add credence to the idea that what we have discovered are different behavioral equilibria along a spectrum of employment contracts. However, the difference in beliefs might also be the result of differential selection. It might be that pessimistic messengers select into performance pay firms while optimists are attracted to coops, for example. We can use the same method as in the previous section to identify selection and isolate the treatment effect of the employment contract on beliefs about the cooperativeness of one’s co-workers.

We report our estimates in Table 3 which is organized similarly to Table 2. The only difference is that we report Tobit regressions instead of multinomial logits. In our baseline, columns (1) and (1r) indicate a significant effect of performance pay on messenger beliefs. In the full sample (1) performance pay couriers believe that their co-workers are 9% less likely to cooperate as a first-mover (p<0.01). Compared to behavior, it appears that selection might affect beliefs because the coefficient from the restricted sub-sample is twice as big (p<0.01). In general, selection appears to dampen the treatment effect of performance pay on beliefs and our interpretation is that the people who are stuck in performance pay firms are particularly pessimistic compared to those who can (or have) switch(ed).

Beliefs, like social preferences, also do not appear to change gradually. In columns (2) and (2r) we add tenure and its interaction with performance pay and find no effect. However, as is the case throughout Table 3, the estimate of the treatment effect in the absence of selection is twice as large. Lastly, in columns (3) and (3r) we see that our point estimates are robust to the addition of demographics. In this case none of the demographics are robust, although they do jointly add to the estimates.

9 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

---

7 OLS estimates with robust standard errors yield the same results.
8 One can also see that the coefficient on performance pay loses significance in (2) but the F-test test of commission and its interaction reveals that they are jointly significant.
social dilemma incentives to see if social preferences are endogenous to the manner of compensation. 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 that they are considerably more cooperative than students in our sequential prisoner’s dilemma experiment. As in Carpenter et al. (2005), this suggests that the workplace frame matters, although comparison of our results with those of 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 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. Third, and perhaps most interestingly, we find 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, 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.

Our results have obvious implications for human relations in the workplace and, potentially, for firm profitability. Rotemberg (1994) demonstrated that altruism could evolve in the workplace if there are strategic complementarities among workers as is the case with our messengers. To some degree our field experiment confirms these predictions because the amount of prosocial behavior that we see is correlated with the pay scheme. We do not, however, find evidence of the sort of enlightened self-interest that is implied by Rotemberg’s model (pp. 686). That is, we see no evidence that our messengers consciously change their preferences. At this point we can only speculate about the implications of social preferences for the overall profitability of firms. While the empirical evidence offered in Lazear (2000) unambiguously suggests that performance pay increases profits when individual effort is identifiable and strategic

---

9 However, we are currently collecting revenue and productivity data from a few of the participating firms. For related work in which a link is made between social preferences and productivity see Carpenter and Seki (2005b).
complementarities are weak among workers, it is yet unknown what happens when complementarities are more important and social preferences are present. In our case, it could easily be that because performance pay “spoils” the pool of social preferences, when commission pay is introduced between-worker cooperation may fall compared to revenue sharing cooperatives, for example, and if the complementarities in the production process are important enough, this could easily result in lower instead of higher profits.
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:

<table>
<thead>
<tr>
<th>Messenger A’s options</th>
<th>Messenger B’s options</th>
<th>Result</th>
<th>Result</th>
<th>Result</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Messenger A cooperates</strong></td>
<td>Messenger B responds by cooperating</td>
<td>$30</td>
<td>$30</td>
<td>$45</td>
<td>$45</td>
</tr>
<tr>
<td><strong>Messenger A does not cooperate</strong></td>
<td>Messenger B responds by not cooperating</td>
<td>$10</td>
<td>$10</td>
<td>$15</td>
<td>$15</td>
</tr>
</tbody>
</table>
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)

- O Cooperate
- O Not cooperate

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

<table>
<thead>
<tr>
<th>If messenger A cooperates, I will</th>
<th>If messenger A does not cooperate, I will</th>
</tr>
</thead>
<tbody>
<tr>
<td>○ Cooperate</td>
<td>○ Cooperate</td>
</tr>
</tbody>
</table>

Please double-check your choices on this sheet.

Please notify the experimenters if you are done with the decision sheet.
11 Tables and Figures

<table>
<thead>
<tr>
<th>TABLE 1: Participant (messenger) demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beliefs: Fraction(As Cooperate)</td>
</tr>
<tr>
<td>Paid on Commission</td>
</tr>
<tr>
<td>Dispatched by Free Call</td>
</tr>
<tr>
<td>Tenant: months at current job</td>
</tr>
<tr>
<td>Hours per week worked</td>
</tr>
<tr>
<td>Household Income (in dollars)</td>
</tr>
<tr>
<td>Swiss</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>Married</td>
</tr>
<tr>
<td>Student</td>
</tr>
<tr>
<td>Years of Schooling</td>
</tr>
</tbody>
</table>
### TABLE 2: The Effect of Performance Pay on Social Preference
( Dependent variable is behavioral type )

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(1r)</th>
<th>(2)</th>
<th>(2r)</th>
<th>(3)</th>
<th>(3r)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prob(Egoist)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paid on Commission</td>
<td>0.125</td>
<td>0.115</td>
<td>0.122</td>
<td>0.143</td>
<td>0.120</td>
<td>0.150</td>
</tr>
<tr>
<td></td>
<td>(0.041)***</td>
<td>(0.055)**</td>
<td>(0.068)*</td>
<td>(0.047)***</td>
<td>(0.066)*</td>
<td>(0.048)***</td>
</tr>
<tr>
<td>Tenure</td>
<td>-0.0003</td>
<td>0.001</td>
<td>-0.0003</td>
<td>0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.001)</td>
<td>(0.003)</td>
<td>(0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tenure × Commission</td>
<td>0.0003</td>
<td>-0.001</td>
<td>0.0002</td>
<td>-0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.001)</td>
<td>(0.003)</td>
<td>(0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Prob(Altruist)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paid on Commission</td>
<td>-0.044</td>
<td>-0.020</td>
<td>0.035</td>
<td>0.026</td>
<td>-0.011</td>
<td>-0.014</td>
</tr>
<tr>
<td></td>
<td>(0.081)</td>
<td>(0.111)</td>
<td>(0.110)</td>
<td>(0.164)</td>
<td>(0.124)</td>
<td>(0.179)</td>
</tr>
<tr>
<td>Tenure</td>
<td>0.0005</td>
<td>-0.0002</td>
<td>-0.001</td>
<td>-0.003</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tenure × Commission</td>
<td>-0.002</td>
<td>-0.001</td>
<td>-0.001</td>
<td>0.0003</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.003)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demographics Included</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>243</td>
<td>153</td>
<td>243</td>
<td>153</td>
<td>235</td>
<td>147</td>
</tr>
<tr>
<td>Pseudo R²</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>0.01</td>
<td>0.06</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Notes: Results are marginal effects from multinomial logit regressions using robust standard errors. The baseline category is conditional cooperator. * indicates significance at the 10%, ** 5%, and *** 1% level.
### TABLE 3: The Effect of Performance Pay on Beliefs

(Dependent variable is beliefs about the fraction of first-movers who cooperate)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(1r)</th>
<th>(2)</th>
<th>(2r)</th>
<th>(3)</th>
<th>(3r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paid on Commission</td>
<td>-0.092</td>
<td>-0.184</td>
<td>-0.083</td>
<td>-0.220</td>
<td>-0.103</td>
<td>-0.239</td>
</tr>
<tr>
<td></td>
<td>(0.036)***</td>
<td>(0.057)***</td>
<td>(0.063)</td>
<td>(0.095)**</td>
<td>(0.064)*</td>
<td>(0.097)**</td>
</tr>
<tr>
<td>Tenure</td>
<td>0.001</td>
<td>-0.001</td>
<td>0.001</td>
<td>-0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tenure × Commission</td>
<td>-0.001</td>
<td>0.001</td>
<td>-0.0005</td>
<td>0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.002)</td>
<td>(0.001)</td>
<td>(0.002)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demographics Included</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>252</td>
<td>157</td>
<td>252</td>
<td>157</td>
<td>244</td>
<td>151</td>
</tr>
<tr>
<td>Truncated Observations</td>
<td>26</td>
<td>15</td>
<td>26</td>
<td>15</td>
<td>26</td>
<td>15</td>
</tr>
<tr>
<td>Pseudo R²</td>
<td>0.09</td>
<td>0.27</td>
<td>0.10</td>
<td>0.27</td>
<td>0.25</td>
<td>0.51</td>
</tr>
</tbody>
</table>

Notes: Results are from tobit regressions. * indicates significance at the 10%, ** 5%, and *** 1% level.
Figure 1: The distribution of types in the sequential prisoner’s dilemma by messenger compensation scheme (Note: No Commission Pay combines hourly paid and revenue sharing messengers).
Figure 2: The distribution of beliefs about how many first-movers will cooperate in the sequential prisoner’s dilemma by messenger compensation scheme (Note: No Commission Pay combines hourly paid and revenue sharing messengers).
Figure 3: The distribution of types by field or lab implementation.
**Figure 4:** The distribution of beliefs about how many first-movers will cooperate in the sequential prisoner’s dilemma by field or lab implementation.
12 References


