IZA DP No. 2400

Incentives and the Allocation of Effort Over Time: The Joint Role of Affective and Cognitive Decision Making

Lorenz Goette
David Huffman

October 2006

Incentives and the Allocation of Effort Over Time: The Joint Role of Affective and Cognitive Decision Making

Lorenz Goette
*University of Zurich, CEPR
and IZA Bonn*

David Huffman
IZA Bonn

ABSTRACT

Incentives and the Allocation of Effort Over Time: The Joint Role of Affective and Cognitive Decision Making

We use natural experiments – plausibly exogenous, anticipated increases in the piece rate – to study how effort responds to incentives. Our first finding, like some previous studies, lends little support to the view that incentives increase effort: raising the piece rate has zero effect on total daily effort. Previous studies have speculated that changes in motivation over the course of the workday, caused by the increase in the piece rate, may lead to this result, but have relied on data aggregated to the day. Our data allow us to look within the workday. We find that workers do respond to incentives within the day: they work significantly harder in early hours of work, but significantly less hard later on, with a net effect of zero on total daily effort. We consider different possible explanations for this behavior. The most parsimonious explanation is a model in the spirit of Loewenstein and O'Donoghue (2005), in which a cognitive system, assumed to behave like the standard economic model predicts, is in conflict with the affective system. We review evidence from psychology and neuroscience to argue that the affective system may be strongly influenced by within-day changes in earnings, relative to an earnings goal. The affective system cares less about income once the goal is surpassed, providing an explanation for a drop in effort later in the day, and for the findings of earlier studies.

JEL Classification: J22, J33, D01, B49

Keywords: labor supply, loss aversion, affect, intertemporal substitution

Corresponding author:

David Huffman
IZA
P.O. Box 7240
53072 Bonn
Germany
E-mail: huffman@iza.org

I. INTRODUCTION

Everyday experience suggests that success and failure on the job can trigger intense emotional reactions. However, models of labor supply typically assume that workers decide how hard to work, and how to respond to incentives, based on a purely cognitive calculation regarding (lifetime) utility. This is contrary to growing evidence that many economic decisions are characterized by a tension between an individual's own (long-term) wellbeing, and temptations to act more impulsively in the heat of the moment (Cohen, 2005). Evidence from brain imaging studies points to emotion, or affect, as the source of these short-sighted impulses. Affective regions of the brain are more strongly activated when making choices that involve immediate benefits (McClure et al., 2004), or when rejecting unfair offers in the ultimatum game (Sanfey et al., 2003). Cognitive areas, on the other hand, are relatively more active when individuals make choices in line with the standard model's predictions (i.e., act more patiently, or accept unfair offers). Corroborating evidence comes from patients with lesions in parts of the brain that integrate the input from affective regions; these individuals exhibit behavior that is more in line with the predictions of the standard economic model, e.g., they are not loss averse (Shiv et al., 2005). Further evidence of this dichotomy comes from studies asking subjects to consciously try to suppress affective reactions (e.g., arousal due to sexual movies; see Beauregard et al., 2001). When trying to suppress affect, cognitive areas of people's brains are more strongly activated. Overall, this body of evidence is consistent with an interpretation where a more deliberative (cognitive) system in the brain – thought to correspond approximately to the standard model in economics – struggles to gain control over the affective system, whose objectives are more strongly influenced by visceral influences, emotions, and moods (Loewenstein and O'Donoghue, 2005).

The tension between affect and cognition may play an important role in work settings. For example, in piece-rate occupations, earnings vary strongly across and within days. Figure 1

shows the leading causes of mood among bicycle messengers in San Francisco, all of whom are paid on a piece rate. Daily income is rated as the most important determinant of mood, much more important than other features in this occupation, like, e.g., behavior of the customers, or the dispatcher. Therefore, as earnings accumulate within a workday, this may induce changes in affect. The impact of this affective terrain within the workday on motivation, effort decisions, and the way that workers respond to changes in financial incentives, has been largely unexplored in economics.

The evidence on how effort responds to financial incentives is mixed. Camerer et al. (1997) and Chou (2002) use data from cab drivers, for whom earnings opportunities vary strongly between days. The standard economic model would predict a positive response of hours worked (the best measure of effort in that context) to an increase in daily wages, yet cab driver studies typically find a negative correlation between wages and hours.[3] Camerer et al. (1997) propose that this behavior reflects a change in motivation within the day, triggered by a wage increase. They hypothesize that cab drivers have a daily income target in mind while working (a claim they back up with survey evidence of managers of cab fleets), and that higher wages make it easier to surpass that target. Surpassing the target leads to a drop in the marginal valuation of income later in the day, which rationalizes lower daily effort. The finding and its interpretation are provocative, however there are serious methodological problems with giving a causal interpretation to the negative correlation between wages and hours worked (Farber, 2005; Goette, Huffman, and Fehr, 2004). Fehr and Goette (2006) randomize commission rate increases

---

[3] Other studies find that increases in the piece rate have weak or only short-lived effects on effort (e.g., Treble, 2003). Paarsch and Shearer (2005) is an exception, finding an increase in effort after an exogenous change in piece rates for tree planters. Evidence from several other studies suggests that switching from fixed wages to piece rates reliably leads to a positive effort response (Lazear, 2000; Shearer, 2004; Goette and Lienhard, 2006). Switches from fixed pay to piece rates, while related, are not the focus of this paper.

among bike messengers in a way that avoids the problem of the cab driver studies. They, too, find a decrease in daily effort while the messengers are on the higher commission rate. The finding is consistent with the original income target interpretation, but because messengers are also observed to work more days during the four weeks of the commission rate increase, it could also be that working more days leads to fatigue, which makes it optimal to reduce within-day effort. A notable feature of all of these studies is that they appeal to a change motivation within a workday to explain the pathological response of total daily effort to incentives, but the data available is aggregated to the day, preventing such changes in motivation from being directly observable.

There is an old literature that does look within the workday, collecting productivity data during different hours of the work. In these studies effort is observed to change over the day in ways that a standard economic model has difficulty accommodating (see the discussion in Oi, 2000). For example, effort profiles in occupations with piece-rate pay show a tendency for output to initially increase, and then decrease later in the day, despite constant economic returns to exert effort during the day (Vernon, 1921; Goldmark and Hopkins, 1920). However, to our knowledge there is no data on how an exogenous change in piece rates affects the motivation to exert effort at different times within a day.

The contribution of this paper is twofold. First, we provide new empirical evidence on how incentives affect effort. At two firms, we observe quasi-experimental changes in piece rates. We also have hourly measures of effort for each worker on each day over several years. This data is ideal for studying whether a change in incentives affects total daily effort, but also how the change affects the motivation to exert effort at different times within the day. We compare within-day effort profiles before and after a messenger receives the increase in the piece rate. For each of the two firms we have also identified a "control" firm, which is very similar except that the piece rate does not change. We can then use differences-in-differences to

identify the impact of the piece rate change on effort profiles, which makes the analysis immune to various criticisms.

The second contribution is a discussion of what kind of model could explain the particular response to incentives that we observe. Understanding the underlying mechanisms is an important step towards understanding what, if anything, can be done to improve the effectiveness of incentives. We focus on two candidate explanations: fatigue and affect. On the one hand, intuition suggests that fatigue could lead to non-standard effort profiles over the day, and hence "pathological" responses to incentives. Previously there has been little in the way of a formal treatment of fatigue, so we propose various candidate models that are consistent with sports physiology. On the other hand, a wage increase affects the accumulation of earnings during a day, and thus could have an impact on motivation to work through the channel of changing affect. One possibility is that affective motivations depend on progress towards a daily income target as hypothesized by Camerer et al. (1997). We consider whether there is more solid basis in evidence from psychology and cognitive neuroscience to construct a model incorporating affect and income targets, and ask whether this can explain the behavior we observe in our data.

The first portion of the paper presents our empirical findings. Our data are the delivery records from four bicycle messenger firms. Bicycle messengers are paid a piece rate, and are ideal for studying the response of effort to incentives as they have relatively large discretion over adjusting their effort within the day. At Firms A and B we observe useful natural experiments: anticipated increases in the piece rate, which occur after a messenger has worked at the firm for several months (the managements use this deferred compensation as a strategy to reduce turnover). We use the third and fourth firms, Firm C and Firm D, to perform even more stringent difference-in-difference tests of the impact of incentives on effort profiles.

We find that the increase in the piece rate has zero effect on total daily effort at Firms A and B, consistent with previous evidence on the weak effect of incentives on daily effort. However, this conceals a strong response in terms of the way that effort is allocated over time within the day. After the increase in the piece rate, our difference-in-difference estimates show that messengers work significantly harder earlier in the day, but significantly less hard later in the day, leading to a net effect of zero on total daily effort. Our empirical strategy is closest in spirit to Farber (2005), who has a similar strategy of looking within the workday. He observes the number of fares taken by a cab driver in each hour of the day, and tests whether accumulating daily income affects motivation to continue working, i.e., the probability that a driver quits in a given hour. Our identification strategy is different, because we study plausibly exogenous changes in the piece rate rather than variation in wages across days. Also, we do not focus solely on quitting, but more broadly on how the change in incentives affects the level of effort in a given hour.[4]

In the second portion of the paper we take our findings further, asking *why* workers re-distribute effort in the way that they do. Our strategy is to consider different possible models of labor supply, and assess which is best able to explain the particular pattern of reallocation that we observe. We first consider the standard model, applied to the decision of how to allocate effort over a workday composed of a series of work episodes. A simple version of the model, with time-separable utility, predicts a non-negative response of effort to an increase in the piece rate during all work episodes. This is inconsistent with the drop in effort at the end of the day observed at Firms A and B.

---

[4] Our data also have the advantage that we can compare the effort of all workers in a given hour on a given day. Thus, we difference out firm*day*hour specific shocks, which is not possible in Farber (2005), because the data typically do not include more than one worker on a given day.

For modeling within-day effort decisions, however, it is more reasonable to allow for non-separable utility, in the form of fatigue spillovers between work episodes. In this case it is less obvious how changing incentives affects the allocation of effort over time. We therefore study models incorporating various forms of fatigue. It turns out, however, that fatigue effects do not overturn the previous result, and that the model still cannot predict a decrease in effort following an increase in productivity.

We then consider a model that captures a similar intuition to Camerer et al. (1997), namely that workers on a piece rate are psychologically motivated to reach a daily income target, or goal, on top of their purely financial motives to exert effort. We provide more solid psychological underpinnings for this hypothesis. Recent evidence from psychology points to affect as the source of the psychological motivation to work towards a goal, and evidence from neuroscience studies show that the involvement of brain regions processing affect are involved in generating this kind of behavior. These stylized facts provide some guidance in formalizing the role of affect in labor supply. We model the worker's effort decision as arising from the interaction of two neural systems, one cognitive, valuing costs and benefits exactly as the standard model assumes, and one affective, exhibiting a valuation of income that depends on an income target. This leads to potentially conflicting motivations. The key feature of the model is that the cognitive system can to some extent control the impulse of the affective system. Nevertheless, the model can explain the types of deviations from optimal effort choices that we observe in our data. We show that for reasonable parameters, a model incorporating an affective system that cares about an income target produces a positive response of effort to piece rates early in the day, and a negative response towards the end of the workday.

The remainder of the paper is organized as follows. In sections II and IIII we describe the empirical setup and our estimation strategy. Section IV presents our empirical results. Section V discusses candidate explanations for our findings. Section VI concludes.

II. EMPIRICAL SETUP

A. Bicycle messenger firms

Bicycle messenger firms offer same-day, or same-hour delivery of packages in most cities around the world, for businesses that require frequent deliveries, e.g., law firms, advertising agencies, and scientific laboratories. A messenger's compensation is typically based on a commission rate, which is a fixed fraction of the price of each delivery and thus is equivalent to a piece rate. Messenger firms price deliveries based on the distance between the pick-up and drop-off points, and the speed with which the customer needs the delivery accomplished.

Bicycle messengers are attractive subjects for studying the impact of incentives of effort because they have substantial discretion over their choice of effort during the day. Deliveries are announced over the airwaves by a dispatcher, and are heard by all of the company's messengers who are working that day. Messengers have several ways to vary effort: they can work hard to finish deliveries quickly, and lobby the dispatcher for more deliveries, or they can make deliveries slowly, and respond slowly to the dispatcher's calls on the radio.

B. Data

The data we use are the electronic delivery records of four different bicycle messenger firms, created by dispatchers at these firms during the course of the workday. The records include information on the price and timing of each delivery made by each messenger on a given day. One of the firms is located in Basel, Switzerland (Firm A), one is in Zurich, Switzerland (Firm C), and the other are in San Francisco, California (Firm B and Firm D). The records cover the following time periods: June 1998 to July 2003 for Firm A; June, 2001 to May, 2003 for Firm B; July 2000 to July 2001 for Firm C; January 1999 to September 2002 for Firm D.

C. Natural Experiments and Control Groups

On the day that a messenger is hired, Firms A and B promise the messenger that his commission rate will increase after he has worked for several months. Firm A raises the commission rate from 38% to 44%, and Firm B raises the commission rate from 50% to 55%. This deferred compensation is intended to give the messenger an incentive to stay at the firm, and thus reduce turnover. Importantly, this strategy requires that the firms make a messenger fully aware of the amount and timing of the future increase in the piece rate. At Firm A the increase must occur at the beginning of a month, due to a monthly pay cycle, and at Firm B the increase is tied to a bi-monthly pay cycle. These institutions create variation in the timing of the commission rate increase across messengers, for example because one messenger starts work on the first of the month, and another starts in the middle of the month. At Firm A the average time a messenger works before the piece rate is increased is 12 weeks, and at Firm B the average is 14 weeks.

As is shown in the next subsection, Firms C and D are very similar to their counterpart in Switzerland or the U.S., respectively, except that they do not change a messenger's commission rate. Throughout their tenure at Firm B, messengers are paid a commission rate of 47%, and messengers at Firm D are paid 50% irrespective of tenure. Thus, Firms C and D have the potential to serve as "control firms" for Firms A and B, allowing us to difference out other factors that might lead to a change in the effort profile over time besides a change in the commission rate. For example, Firms A and C hire only rookie messengers, with the result that their workers experience large productivity gains as they get more experienced. Higher productivity may raise the marginal product of effort, and thus have an impact on effort profiles similar to that of an increase in the commission rate. If this learning on the job takes place over time at Firm A, it could thus lead us to overstate the impact of the change in the commission rate on effort at Firm A. However, since Firm C is very similar to Firm A, our strategy is to use the change in the effort profile at Firm C to difference out this effect. We estimate the change in the

effort profile at Firm C by imposing the same rule that Firm A uses to trigger the commission rate increase, comparing effort profiles before and after 12 weeks of tenure at Firm C. Both Firm B and Firm D typically hire messengers with substantial previous experience, so there is less scope for changes in productivity over time at these firms. In fact, the experience profile is even negative at Firms B and D, although this is difficult to interpret because, e.g., two messengers who both have one day of experience at the firm may differ substantially different in terms of the amount of previous experience at other firms. However, we use Firm D as a control group for Firm B, in order to difference out any other possible factors that might affect effort profiles and be correlated with tenure. When we analyze the data for Firm D, we impose the same 14-week cutoff that Firm B uses for the commission rate increase.

Of course we do not have treatment and control firms in the strictest sense, where workers are randomly assigned to workplaces by an experimenter. Thus, in principle it is possible that different types of workers could self-select into our firms with and without deferred compensation. This is unlikely to be a factor in the case of Firms A and C, however, as these are in different cities that are a considerable distance apart. At Firms B and D, which are both in San Francisco, it seems that selection would mainly take the form of workers preferring, all else equal, to work at Firm B, where commission rates are higher overall. Ultimately, selection would have to take a very special form, in order to cause a problem for our analysis. It is hard to construct a story where selection would bias our result of interest, because our identification strategy is based on comparing the behavior of the same individual over time.

One difference, in terms of the data that we have for the four firms, is the number of messengers who switch piece rates during the sample periods. In the data for Firm A, we observe 110 out of 260 messengers switching from the low piece rate to the high piece rate, whereas in the data for Firm B we only observe 8 out of 25 switching. At Firm C, we observe 51 out of 121 messengers switching from the before-12-weeks to after-12-weeks category, and at

Firm C, we observe 9 out of 25 switching. This means that in the empirical analysis below, the estimates for Firm B (D) will be identified using substantially fewer observations than at Firm A (C), and thus are likely to be estimated less precisely.

<center>

D. Descriptive Statistics

</center>

Overall, the four firms in our study are quite similar, even across countries. They all share the institutional ingredients that are common to bicycle messenger firms, such as commission rate pay, and relative freedom to choose effort during the day. There is a notable difference between Swiss and U.S. firms in terms of the length of the working day, which we discuss below, but within countries firms are similar along this dimension as well. This section provides descriptive statistics showing how work is organized at the four firms, and giving a sense for the comparability of revenues and earnings across firms.

At the Swiss firms, the workday is organized in shifts. A messenger works either a morning shift or an afternoon shift. At the U.S. firms, a messenger works for a whole day. Figure 2 shows distributions of starting and quitting hours. Consider first the Swiss firms, Firms A and C: the morning shift goes from 7:00 or 8:00 in the morning until 12:00 or 13:00 in the afternoon, and the later shift starts at 12:00 or 13:00 and goes until 17:00 or 18:00. A qualitative glance at the histograms suggests that the distributions of start times and quit times of the shifts are very similar across firms. For example, approximately 60% of the morning shifts start between 7am and 8am at both Swiss firms. The other 40% divide in similar proportions between earlier and later start times. Again, the end times of the shifts are similar at both firms: more than 60% of the morning shifts end between 12pm and 1pm at both firms. As the figure shows, the distribution of start times and end times of the afternoon shifts also display the same characteristics. By contrast, the typical workday at Firms B and D begins at 7:00 or 8:00 in the

morning and lasts until 17:00 or 18:00.[5] Again, the times at which the shifts start at the two companies are almost the same, as are the end times.

Table 1 compares labor supply across firms. One measure is hours on duty, defined as hours between the first delivery and the last, including hours during which the messenger had no deliveries (breaks). An alternative measure is active hours, i.e. only those hours in which the messenger had positive earnings, which comes closer to capturing the time spent actively working. At the Swiss firms, most of work time is active. On a typical shift, a messenger has 5 hours on-duty, and 5 active hours. A substantial number of shifts last longer or shorter than 5 hours, however. Notably, the distribution of shift lengths is quite similar across the Swiss firms, for both measures. For example, median work hours and median active hours are the same at both firms. At the firms in the U.S., most workdays involve being on-duty for at least 10 hours, but involve only 8 or 9 active hours. The distribution of hours worked is also comparable across the two U.S. firms, displaying the same median work hours and median active hours.

Within countries, the firms are similar in terms of revenues and earnings. At Firms A and C, average revenues per shift are CHF 224 (USD 179) and CHF 216 (USD 172), respectively. The implied (gross) hourly earnings are CHF 24 (USD 19.2; at the 44% commission rate) and CHF 23 (USD 18.4). At the U.S. firms, average revenues are USD 296 at Firm B, and USD 304 at Firm D. This translates into gross hourly earnings USD 15 at Firm B, and USD 14 at Firm D. Figure 3 provides more detailed picture, showing each firm's revenue profile over the day. The aggregate revenue profile has a similar shape across all firms: an "M-shape", with the dip during the lunch hour. At the two Swiss firms, the first peak occurs slightly later in the morning than at U.S. firms. Otherwise, the profiles of revenues over the day are quite similar across countries,

---

[5] The different organization of the workday partly reflects differences in the workforce: messengers at Swiss firms are often students working part time, whereas messengers in the U.S. are full-time workers.

11

and are even more similar for firms within the same country, both in terms of the level and shape.

III. EMPIRICAL STRATEGY

Our goal is to measure the impact of an increase in the commission rate on total daily effort, and on the way that effort is allocated over the day. We proceed in two steps. First, we estimate the change in daily labor supply that occurred after the change in the commission rate at Firm A and B, using regressions that control for various confounding factors. We discuss this estimation in more detail below. Second, we use the data from Firms C and D to conduct an even more stringent, difference-in-difference test of how the change in the commission rate affected labor supply.

Step One: Our measure of labor supply is defined as follows: We follow each messenger working at Firm A on a particular day for seven hours (where the maximum workday is 7 hours long) and follow each messenger for 12 hours at Firm B (where the maximum is 12 hours). We define hourly labor supply as hourly revenues. This creates seven measurements of hourly labor supply for each workday at Firm A, and twelve for each day at Firm B. If a messenger had zero revenues during a particular hour, we set labor supply to zero in that episode. This measure of labor supply is the broadest possible, and is precisely as standard theory suggests it should be. It captures (i) how hard a messenger is working, (ii) whether he is taking breaks during the day, and (iii) when the messenger quits for the day (after the messenger quits, we set labor supply to zero for the remaining hours in the workday).

We estimate equations of the form

$$rev_{iht} = \gamma^1 high^1_{iht} + \gamma^2 high^2_{iht} \; .. \; + \gamma^m high^m_{iht} + \; x_{iht}b + a_i + d_{ht} + u_{iht}$$

12

where *rev* is hourly revenues of messenger i at hour h on date t (as defined above). Our coefficients of interest are the γ^k's. The variable $high^k$ is a dummy variable equal to 1 if it is the messenger's kth hour of work and he is on the high commission rate, and zero otherwise. We want the γ^k coefficients to reflect the messengers' change in labor supply in work hour k due to the increase in the commission rate. For the coefficients to have this interpretation, we need to control for other factors, correlated with work hour and the commission rate, that affect labor supply.

The vector x contains time-varying individual control variables. Its most important subset consists of dummy variables controlling for the number of hours worked so far. Thus, the baseline effort profile does not affect the estimates of γ. The x vector also contains a set of dummy variables controlling for experience. Especially at Firm A (and Firm C), workers entering the firm are "rookies" with no prior experience as messengers. Hence, there could be significant learning as they acquire more skills during their employment. Because there is variation in the timing of the commission rate increase, we are able to separate learning from the change in incentives. We also include a messenger fixed effect a_i to avoid bias arising from the possibility that below-average messengers drop out before they get promoted. Finally, we include an hourly fixed effect d_{ht}, which we estimate separately for each hour on each day to control for the time profile of the availability of deliveries.

With these controls in place, γ^k indicates by how much the messenger changed labor supply in work hour k following the increase in the commission rate. We estimated the equation above using OLS. An important issue is how one should calculate the standard errors of the estimated coefficients. Given the hourly frequency of our measures, there are various ways in which u_{it}, the error term, departs from the i.i.d. assumption of OLS. First, the way we construct

our measure of labor supply makes the error term inherently heteroskedastic.[6] We correct for this by estimating robust standard errors. Second, there are two potential sources of correlation between the error terms. Within a given day, if one messenger was assigned a delivery, another messenger will end up with one less delivery. This leads to negative correlation of the residuals within a day, rendering OLS standard errors too large. On the other hand, there could be positive correlation in u_{it} for observations coming from a given messenger, rendering OLS standard errors too small (see Bertrand, Duflo and Mullainathan, 2004, for an extensive discussion). We are left with no alternative but to report two sets of standard errors. One set is adjusted for clustering on days. Because this ignores the (potentially) positive correlation within individuals, we consider these standard errors the lower bounds. The other is adjusted for clustering on messengers. We consider this the upper bound on the standard errors, because it ignores the (potentially) negative correlation within days. However, most of our conclusions do not depend on which adjustment of standard errors we use, and if they do, we point it out in the discussion.

Step Two: Our most conservative estimate is obtained by contrasting the estimates from the firms where a commission rate change occurred to estimates from the firms where no change occurred. For the control firms, we apply the same tenure cutoffs as used at the treated firms, and run the same first-stage regression in order to estimate the impact of a "placebo intervention." More formally, denote the estimate of the change in labor supply at a treated firm by γ^{treat}, and the covariance matrix by Σ^{treat}. The corresponding estimates for the control firm are denoted $\gamma^{control}$ and $\Sigma^{control}$, respectively. Then, the difference-in-difference estimate of the impact of the treatment is $(\gamma^{treat} - \gamma^{control})$, with covariance matrix $\Sigma^{treat} + \Sigma^{control}$. Under the null

---

[6] Because our dependent variable is bounded by zero from below, this necessarily implies that the variance of the error term differs between observations.

of no treatment effect, $(\gamma^{treat} - \gamma^{control})'(\Sigma^{treat} + \Sigma^{control})^{-1}(\gamma^{treat} - \gamma^{control})$ has an asymptotic Chi-square distribution with m degrees of freedom.[7]

IV. RESULTS

A. Total Daily Labor Supply after the Change in the Commission Rate

We begin by comparing our results to earlier studies, which have primarily relied on labor supply aggregated to the day. For ease of comparison, the first row of Table 2a shows the impact of the increase in the commission rate at Firms A and B on total daily labor supply, i.e., the sum over all estimated γ^k's. In the first columns for Firms A and B, we show that higher commission rates are associated with higher total revenues per day, although the effect is not robust to clustering on messengers. But since we have not yet controlled for messenger-specific differences, this should not be interpreted as the change in behavior due to the higher commission rate. In fact, the second columns for Firms A and B show that when we include messenger fixed effects, the relationship between higher commission rates and total revenues vanishes. At Firm A, the point estimate is CHF 1.58, which is tiny relative to average daily revenues of CHF 250. The point estimate at Firm B is actually negative, although it is not significant at the five percent level. Again, the point estimate is quantitatively small (USD 8.25 relative to average revenues of USD 290 per day). Table 2b shows the results for Firms C and D, before and after the artificially imposed cutoffs of 12 and 14 weeks respectively. As expected, we do not find a significant change in total daily labor supply after these imposed cutoffs.

---

[7] This is equivalent to estimating a fully interacted model for the two firms and testing for significant interaction effects in the change of the revenue profile between the treated and control messengers.

In summary, the results in Table 2a and Tb are broadly consistent with the findings from other studies that find only a weak effect of a change in piece rates on total daily effort. The increase in the commission rate seems to have no labor supply.

B. The Effort Profile after the Change in the Commission Rate

Based only on our findings at the level of daily labor supply, one might conclude that an increase in the piece rate has little or no impact on effort. However, we are able to go beyond previous studies, and compare within-day effort profiles before and after an increase in the commission rate. The evidence we find lends no support to the hypothesis that effort is unresponsive to changes in incentives, and shows that analysis based only on daily aggregates can be misleading.

Returning to Table 2a, the second row shows that the hypothesis that all γ^k are equal to zero is overwhelmingly rejected in all cases. This is also shown graphically, in Panels A and B of Figure 4, where the γ^k coefficients are plotted over time, with error bars for plus and minus one standard error of the estimate. These graphs show a strong response to changes in the commission rate. At Firm A, there is a distinct pattern of initially increasing and then decreasing effort over time, in response to the increase in the commission rate. Because the large increase in effort early in the day is offset by an almost equally large decrease of effort towards the end of the day, the result is a tiny net change in total labor supply for the day. The distortions in labor supply throughout the day, however, are quantitatively significant. Average revenues are roughly CHF 50 in each hour at Firm A if a messenger is active, implying that the elasticity of revenues with respect to the commission rate is greater than 1 early in the day, but becomes almost -1 one later in the day. At Firm B, there is a similar time profile over the day. There is a clear decline in the effort profile after the increase in the commission rate, resulting in a statistically significant decrease in effort levels towards the end of the day, with negative

elasticities on the same order of magnitude as those observed at Firm A. There is somewhat less evidence that the messengers exert more effort early in the day, because the coefficients are estimated less precisely: a test of the hypothesis that the first four coefficients add up to zero at Firm B is inconclusive ($p < 0.05$ when clustering on days, but the effect is not significant when we cluster on messengers). However, we can easily reject the hypothesis that the coefficients for the last four hours of work add up to zero, at the 1% significance level.

For completeness, in Figure 6 we also report the baseline effort profiles, before workers experience the experimental variation in piece rates, or placebo interventions, respectively. At all firms the profiles are downward sloping, exhibiting the sharpest and most significant drop in the latter portion of the day. The shape of the profiles is unsurprising, mainly because there is a strong tendency to find a downward sloping profile, for mechanical reasons. As workers begin to quit, they start to appear in our data as zeros, thereby leading to a drop in revenues.

C. Difference-in-Difference Estimates

If messengers learn on the job at Firms A and C, this increase in productivity raises the marginal product of effort, and thus may act like a commission rate increase. As a result, the estimate we get for the effect of the commission rate increase at Firm A may be biased upward, because it partly reflects this increase in productivity that is also correlated with tenure. Figure 6 displays the estimated experience profile at Firms A and C. Both profiles show clear increases in the productivity of messengers as experience increases.[8] These gains are slightly more pronounced at Firm C, but are of a similar order of magnitude. For example, relative to a rookie, 12 work weeks of experience raise revenues per hour by CHF 10, almost 20 percent, at Firm C, and CHF 8 at Firm A.

_____

[8] As discussed above, we find no indication of productivity gains on the job at Firms B and D, consistent with these firms' policies of hiring only experienced workers.

Panel C of Figure 4 shows the estimates of the change in the revenue profile that occurs at Firm C, imposing the same cutoff used by Firm A to trigger the increase in the commission rate. We see a response that is much less pronounced than, but qualitatively similar to the change observed at Firm A after the commission rate increase. Even though the individual standard errors are quite large, a statistical test clearly rejects the hypothesis that all coefficients are equal to zero at Firm C (Table 2b). This could reflect the similarity across the two firms in terms of an underlying process of learning on the job. By contrast, when we perform the analogous exercise at Firm D, we do not find the same downward-sloping pattern observed at the other three firms. Although statistical tests reject that all coefficients are zero, there is no clear pattern in how the effort profile changes. Thus, while the rookies at Firm C eventually exhibit a pattern similar to that observed after the wage increase at Firms A and B, there is no such pattern evident for the experienced messengers at Firm D. This is consistent with the interpretation that the gains in productivity of the messengers at Firm C act somewhat like an increase in the commission rate, and thus have a similar although weaker impact on the shape of the effort profile. The messengers at Firms B and D, however, are experienced, and thus there is probably much less scope for productivity gains over time. The fact that there no gains to experience at Firm B or D strengthens this interpretation.

For our most conservative estimates of the impact of the change in the commission rate, we difference out the changes in the revenue profile observed at the respective control firms. Panels E and F display these results. The pattern remains downward sloping, and the order of magnitude of the change in labor supply in a given hour is still considerable at both firms. The hypothesis that all coefficients of the difference-in-difference estimate are equal to zero is clearly rejected in both cases. Thus, even in our strictest test, there is clear evidence that the revenue profile changes in response to a commission rate increase in a very similar way at Firms A and B.

In summary, our empirical analysis reveals a strong within-day impact of financial incentives. Workers increase effort early in the day, but this is offset by a strong reduction in effort later in the day. Evidence that a wage increase can change the way that workers allocate effort within the workday is important in its own right, but the particular, offsetting reallocation that we observe also helps explain why effort may appear unresponsive to incentives in data that are aggregated to the day.

V. DISCUSSION

In this section we take our results further, and discuss *why* workers might reallocate effort over time after a wage increase, in the particular way that they do. Understanding the source of this behavior is clearly important for understanding what, if anything, can be done to improve the effectiveness of incentives. We contrast the predictions of three types of models. The first is the standard model of labor supply, applied to the decision of how much effort to exert in a series of work episodes during a workday. The second model incorporates a realistic form of fatigue, such that effort in one period increases the costs of effort in the next. The third incorporates affect as a source of motivation, causing the worker to have a non-linear valuation of income over the day as he works towards a daily earnings goal.

A. Within-Day Labor Supply and Time-Separable Preferences

The prototypical economic model assumes that individuals maximize a time-separable utility function over their lifetime, or more realistically, over an extended period of time such as several years. In this paper, we are interested foremost in the allocation of effort within a given workday, during such an extended period of time.

To fix notation, divide a workday k into m work episodes of length Δ. We adopt the convention that episode t lasts from t to $t + \Delta$ on day k. We do not need to solve the full

lifetime optimization problem in order to characterize the allocation of effort to different work episodes within a day; intertemporal maximization implies that the optimal choice of effort within each work episode t has an equivalent representation as

$$\underset{e_t}{Max}\; V_t = \lambda w_k e_t - c(e_t).$$

(1)

The parameter λ is the marginal utility of lifetime income at the optimum, w is the (discounted) wage on workday k, and in a slight abuse of notation e_t denotes labor supply in episode t on workday k. $c(\cdot)$ is the convex cost of effort in that work episode.[9] A noteworthy feature of the individual's choice problem is that small changes in w lead to only small changes in lifetime wealth, and thus leave λ essentially unchanged. Hence, the valuation of income in each episode should be linear. We discuss below how the role of affect in decision-making may overturn this prediction and lead to non-linear valuation of income over the course of a day.

The optimality condition for effort in episode t is

$$c'(e_t) = \lambda w_t$$

(2)

This condition summarizes a powerful intuition, which is that a wage increase calls for higher effort (at all points during the workday). This is true for unanticipated wage changes as long as the wage change is small, because a small change in the wage rate leaves lifetime wealth, and λ, essentially unchanged. If the wage increase is anticipated, as is the case at the firms studied in our empirical analysis, the intuition holds for any size wage increase: income effects are entirely absorbed in λ as soon as the individual learns about the wage profile, and thus effort must increase when the wage increase finally occurs (see also Browning, Deaton, and Irish, 1985).

---

[9] The convexity of $c(\cdot)$ follows from the concavity of the underlying intertemporal preferences (see the appendix in Fehr and Goette, 2002, for a proof).

One qualification to this intuition is also apparent in the same condition. The curvature of $c'()$ limits the response of effort to the wage. For sufficiently steep marginal costs, not responding to the higher wage may be optimal. Given that we find no response in total daily labor supply to the increase in the piece rate, this could be consistent with steep marginal cost of effort making non-response optimal. However, in this case, it must be optimal to not respond at all points during the day, and this is inconsistent with the strong within-day response observed in our data.

In summary, the simple time-separable model predicts a positive response of effort to an increase in incentives, in all periods of the day, unless the marginal costs of effort make non-response the optimal choice in all periods. Neither prediction matches the data: we observe a strong response, which includes a substantial decrease in effort over the latter portion of the day.

B. Fatigue

While the time-separable model is useful for illustrating the basic intuition of the standard labor supply framework, it seems important to consider the possibility that effort in one period could raise the costs of effort in the next period. We refer to such non-separabilities in effort costs as fatigue. Little attention has been paid to how fatigue affects the optimal choice of labor supply over the workday, and how fatigue might interact with the effects of a wage increase. In this subsection we consider models incorporating different forms of fatigue.

We first consider a general specification for fatigue that we call "fatigue with recovery", which is flexible, and realistic in the sense that it is consistent with evidence from sports physiology. We assume that effort in one period adds to a stock of fatigue, which raises the marginal costs of effort in the next period. We say that an individual "recovers" from fatigue if, having reduced effort for one episode t, the individual enters the next episode with lower

marginal costs. Formally, effort costs in period t depend on effort e_t and the "stock" of fatigue $k_t = e_{t-1} + \delta\, k_{t-1}$ with $0 < \delta < 1$. We assume that

$$c(e_t, k_t) = c(e_t + \delta \cdot k_t)$$

Thus, the stock of fatigue decays exponentially, capturing the idea of recovery. If an individual rests, this will lower the marginal costs in the next period. In the appendix, we show that the optimal effort profile in this case is u-shaped, for an arbitrary number of work episodes m and for any convex $c(\cdot)$. The intuition behind this profile is straightforward. The marginal benefit of effort is equal to the wage, which is constant across periods. Thus, optimality dictates that the marginal costs of effort are constant across periods as well. Considering the second work episode, a fraction δ of the initial effort level e_1 has decayed, so keeping the marginal cost of effort constant implies setting $e_2 = (1 - \delta)e_1$ in this period. Therefore, effort is higher in the first episode than in the second. The condition that describes second-period effort also applies to all subsequent periods until the last, dictating a constant effort-level during these periods that is proportional to first-period effort. In the last period effort increases, surpassing even first period effort, because at this point there are no future consequences in terms of higher effort costs. Notably, a u-shaped profile is one of the key predictions of "energy flow models" from sports physiology, applied to the problem of optimal pacing for cycling races (de Koning et al., 1999). Several experiments in this literature suggest that individuals are in fact well able to carry out this strategy of optimal pacing (Foster et al, 2003a, 2003b, 2005; de Koning et al., 2005). The profile also matches the strategies for pacing suggested by guides for long-distance runners (e.g., Glover and Glover, 1999).

It turns out that the optimal response of effort to a wage increase is strictly positive in the first period, and in all subsequent periods up until the final period, in the case of fatigue with recovery. Even in the final period, the response of effort must be positive, unless there is

extreme curvature of the *marginal* cost function. In the appendix we show that a drop in effort in the final period occurs only if the curvature of marginal costs in that period are at least 4 times as steep as the curvature in the second-to-last period, i.e., the worker would need to be paid 4 times as much to exert effort in the final work episode as opposed to the second-to-last episode. This is a lower bound, based on assuming the value of δ that minimizes the amount of curvature needed to generate a decrease in effort. Based on these results, we conclude that fatigue is an unlikely to lead to a strong negative response of effort to a wage increase, over a substantial portion of the day, as is observed in our data.

It is interesting to investigate whether the results would change if workers were myopic with respect to fatigue. Although the assumption of myopia implies that the worker does not take into account the impact of current effort on future effort costs, the worker still chooses optimally within each work episode. In the appendix we show that in this case effort in every period, including the last period, is proportional to effort in the first period, irrespective of the shape of the cost function. Effort is highest in the first period, and then constant thereafter, because the myopic worker still sets the marginal cost of effort equal to the wage in each period. In this case, the optimal response of effort to a wage increase is strictly positive in the first period. Thus, it is positive in every subsequent period as well.

One might expect that myopia would have the most extreme consequences given a less realistic form of fatigue where resting does not help, and the fatigue stock simply builds up over the day. In this case the fatigue stock is $k_t = e_{t-1} + k_{t-1}$ and effort costs are given by $c(e_t, k_t) = c(e_t + \beta k_t)$ with $\beta < 1$. The resulting effort profile is decreasing over time, which is more consistent with the baseline effort profiles observed in our empirical application.[10]

---

[10] By contrast, forward looking behavior with this type of fatigue generally implies an upward sloping effort profile. Intuitively, a forward-looking individual finds it optimal to postpone effort

However, as shown in the appendix, in the case of fatigue without recovery effort is again proportional to first-period effort, and the response of effort to a wage increase must be strictly positive in the first and all subsequent periods, for any convex effort cost function.

More fundamentally, it is difficult to explain the behavior we observe with any sort of non-separable effort costs, once we take into account the additional empirical fact that total daily effort remains unchanged in our data, before and after the increase in the piece rate. Holding total daily effort constant, it cannot be optimal to reallocate effort across periods, because this just distorts marginal costs. Put another way, the optimal way not to respond to a wage increase is to not respond at all.

In summary, incorporating fatigue into the standard model of labor supply does little to overturn the intuition that effort responds positively to a wage increase. Effort can decrease only in the final period out of m work episodes, and then only given very extreme curvature in the marginal cost function of effort. Adding the assumption of myopia, the response of effort is unambiguously positively in all periods, including the last. This is true even in the extreme case of fatigue with zero recovery. Finally, taking into account the empirical fact that total daily effort remains constant after the wage increase, it is difficult for any sort of non-separable effort costs to explain a reallocation of effort within the day. We conclude that fatigue is not a plausible explanation for the response in the data.

---

in this case, because effort at the end of the day does not have consequences for future effort costs. See the appendix for further results.

C. Affect as a Source of Motivation

In this section we focus on an explanation in which a change in the wage can change the marginal valuation of income within a workday.[11] Recall that in the introduction we provided evidence that daily earnings are an important determinant of mood for workers in piece-rate occupations. To the extent that a wage change alters the amount of earnings a worker has achieved by a given point in the day, this could also have an impact on affect, and the motivation to work. In this section we follow Loewenstein and O'Donoghue (2005) by assuming that an individual's valuation of income is influenced by both affective and cognitive decision-making systems, and develop a new model of labor supply. The cognitive system is just the standard model of labor supply. Our assumptions regarding how the affective system's valuation of income changes over the day are guided by evidence from psychology and neuroscience.

The role of affect in decision-making: A key finding from neuroscience is that cognition and affect are governed by distinct neural systems in the brain (for an overview see Cohen, 2005), and can generate conflicting motivations. The survival-oriented affective system has a relatively "conservative" set of hardwired priorities, whereas the cognitive system takes into account more global, long-term considerations.

One example of the affective system's conservatism is a tendency to prioritize immediate rewards and threats over longer-term considerations. The affective priority on immediate rewards is evident in the tendency for people to choose a small, immediate reward over a larger, delayed reward if the immediate reward is visible at the time of the decision (Mischel *et al.*, 1972; Mischel *et al.*, 1989; Mischel *et al.*, 2003). McLure et al. (2005) show that choices

---

[11] Note that an ordinary income effect, arising from a permanent wage increase, could explain a reduction in the valuation of income over the rest of the lifetime, but cannot explain why the valuation of income should change systematically over time within a workday.

involving an immediate reward were associated with stronger activation of affective regions in the brain than choices that did not involve an immediate reward.

The affective system is also conservative regarding losses. In choice experiments, people typically evaluate outcomes as gains or losses relative to a reference level, and exhibit *loss aversion*, disliking losses more than they like gains of the same amount (for a review see Tversky and Kahneman, 2000). In a brain lesion study, Shiv *et al.* (2005) provide evidence that loss aversion is generated by the affective system of the brain. Normal patients exhibited loss aversion in a real-stakes lottery task, whereas subjects with damage to the ventro-medial prefrontal cortex, a brain region involved in processing affect, did not. Chen *et al.* (2005) provide further evidence that loss aversion is seated in the evolutionarily older, affective structures of the brain, shared by both humans and other animals, by demonstrating loss aversion among capuchin monkeys.

A number of studies provide direct evidence on the importance of affective reactions for determining task effort. For example, Heath, Larrick and Wu (1999) posed subjects with different hypothetical scenarios in which individuals had done the same number of sit-ups but had different personal goals. Subjects predicted that individuals who were below their goal in these scenarios were experiencing more emotion, and would work harder to perform an additional sit-up. Both findings are consistent with the view that loss aversion, triggered by an affective reaction, serves as a source of motivation to work towards a goal. Subjects were also posed with a different scenario in which two individuals have completed the same number of sit-ups, and are both below their goal, but have different goals. Subjects indicate that the individual with the closer goal will work harder to perform one additional sit-up, consistent with increasing affective arousal as goal becomes more immediate. A pattern of increasing effort as a goal or reward draws is known as the *goal-gradient* effect, and has also been documented many times in

animal studies, e.g., rats run faster in a straight chute when they get closer to a reward (Hull, 1934; for a review see Heilizer, 1977).

More recently, animal studies have provided evidence on the role of the affective system in generating motivation to work towards a goal, at the neurological level. Shidara, Aigner, and Richmond (1998) and Shidara and Richmond (2002) monitored brain activity in monkeys as they worked towards a reward, and found selective response in the ventral striatum and anterior cingulate, respectively, as visual cues signaled decreasing distance from the reward (distance was varied randomly over time, so monkeys had to rely on cues to infer current proximity). These brain structures are believed to play a role in a loop between reward expectancy, affective response, and effort. At the same time that the monkeys exhibited increasing activation in these parts of the affective system, they exhibited a goal gradient, increasing effort and making fewer mistakes on the task as distance to the goal decreased.[12]

Incorporating affect into a model of labor supply: The evidence surveyed above suggests a new model of labor supply, in which a worker's effort decisions during the day reflect both affective and cognitive motivations. Similar to Loewenstein and O'Donoghue (2005) and other "dual-process" models in economics (Thaler and Shefrin, 1981; Shefrin and Thaler, 1988; Metcalfe and Mischel, 1999; Bernheim and Rangel, 2003; 2004; Benhabib and Bisin, 2004; Fudenberg and Levine, 2004), we assume a two-part objective function for the worker. The first captures the preferences of the cognitive system, and values income linearly over the course of work period *t,* exactly as in the standard model of labor supply. More formally, utility in period *t*, from a cognitive perspective, is given by:

$$U_t = w_t e_t - c(e_t) \tag{3}$$

---

[12] For evidence of a similar goal gradient effect among humans, see, e.g., See, Heath and Fox (2003), Kivetz, Urminsky, and Zheng (2005).

Where the marginal utility of lifetime income, λ, is normalized to 1, w_t is the wage in period t, and $c(\cdot)$ is a convex cost function of effort. We denote the cognitive optimum by $e_t^C = \arg\max U_t$.

The second portion of the worker's objective function captures the preferences of the affective system. Consistent with reference-dependence, the valuation of income over the day is assumed to vary with distance from a daily earnings goal, denoted r. The utility of effort in period t for the affective system is given by:

$$v(w_t e_t - r) - c(e_t) \tag{4}$$

We assume that the marginal benefit of income to the affective system, $v'()$, is increasing as total daily earnings approach r from below, reflecting the affective system's sensitivity to immediacy. Once total earnings have surpassed r, however, $v'()$ is assumed to decrease with further earnings, reflecting a dissipation of affective arousal. Furthermore, $v'(-x) > v'(x)$ for all x, i.e., there is a kink in the affective system's value function at the goal, consistent with loss aversion.[13] As a result of the kink, the affective system's marginal valuation of income drops discretely after reaching the goal. We denote the affective optimum by $e_t^A = \arg\max V_t$. This implies that the affective system prefers the effort level that solves

$$c'(e_t) = w v_t' \tag{5}$$

where v_t' reflects the affective system's marginal valuation of income given the decision maker's choice of effort, and the resulting distance from the goal in period t.

---

[13] The function $v()$ is thus equivalent to the value function proposed by Kahneman and Tversky (1979). In this sense our model is similar to Wu, Heath, and Larrick (2002), who propose a dynamic, value-function based model of working towards a goal. An important difference is that they assume the individual is myopic, whereas we maintain the assumption of forward-looking behavior.

We combine the cognitive and affective components into a single objective function, as in Loewenstein and O'Donoghue (2005), and assume that the worker tries to achieve the cognitive optimum, e^C, in each work period, subject to willpower costs involved in moving away from the affective optimum, e^A. For simplicity, we assume that a workday consists of only three work episodes. The individual thus maximizes the following objective function:[14]

$$
\begin{aligned}
\underset{e_1, e_2, e_3}{Max} \ Q_t = {}& w_1 e_1 + w_2 e_2 + w_3 e_3 - c(e_1) - c(e_2) - c(e_3) \\
& - h \Big[v(we_1^A - r) - c(e_1^A) - (v(we_1 - r) - c(e_1)) \Big] \\
& - h \Big[v(w(e_1 + e_2^A) - r) - c(e_2^A) - (v(w(e_1 + e_2) - r) - c(e_2)) \Big] \\
& - h \Big[v(w(e_1 + e_2 + e_3^A) - r) - c(e_3^A) - (v(w(e_1 + e_2 + e_3) - r) - c(e_3)) \Big]
\end{aligned}
\tag{6}
$$

Willpower costs in each episode are captured by the terms in brackets, which express the difference between the affective system's objective function, evaluated at the affective optimum, e^A, and the affective system's objective function evaluated at the worker's chosen effort level. Willpower costs are thus equal to zero if the worker complies with the wishes of the affective system, and increase linearly in deviations from e^A. The optimal effort levels are then given by the following first order conditions:

$$
\begin{aligned}
c'(e_1) &= w \frac{1 + hv_1' - h\big[v_{A,2}' - v_2'\big] - h\big[v_{A,3}' - v_3'\big]}{1 + h} \\
c'(e_2) &= w \frac{1 + hv_2' - h\big[v_{A,3}' - v_3'\big]}{1 + h} \\
c'(e_3) &= w \frac{1 + hv_3'}{1 + h}
\end{aligned}
\tag{7}
$$

Where $v_{A,t}'$ denotes the marginal utility of money to the affective system at the affective optimum, in period t, and v_t' is the affective system's marginal valuation of income given the

---

[14] To be entirely precise, we solve for the perception-perfect strategy (analogous to, e.g., O'Donoghue and Rabin (2000), assuming sophistication. However, since the individual is time-consistent, we can simply apply maximization to (6).

decision maker's actual choice of effort in period t. These conditions have a straightforward interpretation: the individual chooses effort in each episode so that the marginal cost of effort equals the marginal benefits. The marginal benefit depends on the marginal value of money to the cognitive system, w, and on how the chosen level of effort e affects willpower costs. The decision maker takes into account that increasing e by one unit decreases the distance to the affective optimum in the current episode by v'_t, and hence current willpower costs by hv'_t. The additional terms capture the decision maker's consideration of future willpower costs. If current effort tends to raises the difference between the marginal valuations at the future affective optimum and at the actual future effort choice, $v'_{A,s} - v'_s > 0$, this raises future willpower costs and provides a motive for lowering effort in the current period. If, on the other hand, current effort reduces willpower costs in future episodes, this generates an extra incentive to increase effort in the current period.

The model has several new and interesting implications: A first observation is that affect can lead to either lower or higher effort levels, compared to effort levels predicted by the standard model. As can be see from (7), effort is higher in period 3 than in a purely cognitive model if the affective system's marginal valuation of income, v'_3, is greater than 1, the value the cognitive system places on an additional dollar (recall that λ was assumed to be equal to 1). If $v'_3 < 1$, effort is lower than in a purely cognitive model. A second observation is that the effort levels can display the goal gradient effect. For example, suppose it is optimal for the decision maker to be below the goal in the first two periods and above the goal in the final period, taking into account willpower costs. For the sake of the example, assume that the affective system would prefer being above the goal already period 2 (and, hence, also in period 3). Rewrite (7) to obtain

$$c'(e_1) = c'(e_2) + w\frac{h}{1+h}\left[v_1' - v_{A,2}'\right]$$
$$c'(e_2) = c'(e_3) + w\frac{h}{1+h}\left[v_2' - v_{A,3}'\right]$$

(8)

If $v_1' - v_{A,2}' < 0$, there is an increasing effort profile over the first two periods, i.e., $e_2^* > e_1^*$. This will be true as long as the kink in $v()$ at the goal is not too extreme, and $v()$ is sufficiently convex below the reference point. Intuitively, these properties makes it possible for v_1' to be lower than the marginal utility at the affective optimum in period 2, $v_{2,A}'$, even though the affective optimum in period 2 is above the goal and thus is in the region where the $v()$ function is beginning to decrease. Now consider period 3, in which the worker is above the income target. The concavity of $v()$ in this region implies $v_2' - v_{A,3}' > 0$, and thus a drop in effort, i.e., $e_3^* < e_2^*$. Putting these results together, the model generates a pattern of increasing effort leading up to the goal, and a drop in effort after the goal, consistent with the goal gradient.[15] If, on the other hand, the kink at the goal is more extreme, and $v()$ is not too convex below the goal, this yields $v_1' - v_{A,2}' > 0$, which implies $e_3^* < e_2^* < e_1^*$. Therefore, depending on the extent of the kink and curvature in $v()$, the model can predict a goal gradient effect, similar to what is observed in the experiments reviewed above. On the other hand, the model can also generate a strictly

---

[15] The resulting effort profile looks like a "naïf" goal gradient effect, but arises from a more complex calculus: the decision maker takes into account the fact that effort in episode 1 induces the same changes in willpower costs in the future as effort in period 2, but, in addition, also changes the marginal affective valuation of money in episode 1, v_1', and the marginal valuation at the affective system's optimum value in episode 2, $v_{A,2}'$. Thus, the difference $v_1' - v_{A,2}'$ determines relative effort in the two periods. A similar tradeoff determines relative effort in episodes 2 and 3.

downward-sloping effort profile, similar to what we observe in the empirical application in the previous section.[16]

The key question is whether this model can explain the results obtained from the natural experiments observed in our data: an increase in effort early in the workday, but a drop in effort later in the day, following an exogenous change in the piece rate. It turns out that the model can generate this pattern, precisely because a wage increase changes the valuation of income over the day. To see this, consider a simpler version of the model with two-periods. Suppose that it is optimal for the individual to be below the goal for the first period and above the goal in the second, and assume that the affective system prefers to put in just enough effort in the second period so that the goal is achieved, but no more, i.e., $e_2^A = \dfrac{r - e_1 w}{w}$. Note that this is not a particularly restrictive example; there is a whole range of values for e_1 such that the affective system prefers to just reach the kink in the second period, because the marginal valuation of income for the affective system drops discretely at this point. Because in this case e_2^A adjusts to small changes in e_1 in such a way as to keep $v_2^A = v(0) = 0$, changing e_1 has no impact on future willpower costs, through the channel of changing the marginal valuation of income at the affective optimum, $v_{2,A}'$.[17] Thus, terms involving $v_{2,A}'$ drop out, and the first order conditions simplify to:

---

[16] As we note in the previous section, there might be other, mechanical, reasons for this profile to be downward-sloping. Thus, while it is comforting that our model can accommodate this behavior, the empirical finding must still be interpreted with caution.

[17] Note that even though e_1 does not change the affective optimum, it still has an impact on future willpower costs, which the decision maker takes into account. This occurs through the channel of v_2', because working harder changes the distance from the goal in period 2.

$$c_1' = w \frac{1 + hv_1' + hv_2'}{1 + h}$$

$$c_2' = w \frac{1 + hv_2'}{1 + h}$$

From these conditions it is clear that a wage increase causes the cognitive system to place a higher value on effort in the first period, for two reasons: due to the higher financial incentives, and also because the higher wage places the worker closer to the goal in period 1 for any given effort level, thereby increasing the affective system's valuation of effort in period 1 as well.[18] On the other hand, the cognitive system is aware that increasing first-period effort increases the distance from the affective optimum in period 2, thereby increasing future willpower costs. This provides a motive for lowering first-period effort. In Appendix B we derive sufficient conditions for the first two effects to dominate the third, with the result that period 1 effort increases in response to a wage increase, consistent with the behavior we observe in the first portion of the workday.[19] Turning to the second period, it is clear that a wage increase raises the value of effort due to financial incentives, but decreases the affective system's valuation of income because the individual is already above the goal. Because affect only works against a wage increase in the second period, an immediate implication is that the effort response will be smaller in period 2 than in period 1. Furthermore, the conditions derived in the appendix are sufficient for the second effect to dominate the first effect in period 2, thereby leading to a decrease in period 2 effort after a wage increase, consistent with the drop in effort observed in the latter portion of

---

[18] Holding effort constant, the marginal valuation to the affective system, v_1', is increasing in w because v is convex below the earnings goal.

[19] It is clear that similar conditions will exist for a case with more than two periods, but this complicates the comparative statics without adding to the main insight offered by the simpler model.

the workday in our data. Although terms such as willpower costs are difficult to calibrate, the conditions in the appendix hold for a reasonable range of the other parameter values.

In summary, evidence from psychology and neuroscience suggests that affect is an important source of motivation when people work on a task. A model of labor supply that nests the standard, cognitive model, but also incorporates affective reactions triggered by progress towards a daily goal, can explain the response to the wage increase observed in our data.

VI. CONCLUSION

In this paper, we present to our knowledge the first evidence on how exogenous changes in the piece rate affect the allocation of effort over the workday. Our data come from four bike messenger firms that allow us to examine labor supply over time. Two of the firms grant bike messengers an increase in the piece rate after a pre-specified time, while the other two do not. We use the timing of these increases to examine the impact of a change in the piece rate, using the other two firms as controls. Consistent with earlier studies, we find no overall effect of the change in the piece rate. However, we find strong effects of the change in the piece rate on how effort is allocated over the workday. Bike messengers exert more effort early in the workday, but work significantly less hard towards the end of the day. The latter effect is so strong that it completely undoes the increase in labor supply early in the day.

We then discuss two broad classes of models that could potentially explain this result. The first class of models tries to capture the change in the effort profile by assuming that fatigue may develop over the day. Surprisingly, these models are unable to reconcile our findings for at least two reasons. First, there is the inherent difficulty to explain why the allocation of effort is changed at all when overall effort remains constant, which none of these models can explain. Second, even when one assumes that individuals make the systematic mistake of not taking the

intertemporal spillovers of their decisions into account, the model robustly predicts an increases in labor supply at each point in time, as we show.

The second class of models builds on evidence from psychology and neuroscience to specify how the valuation of money may change over the course of a workday. In the spirit of Loewenstein and O'Donoghue and (2005), behavior is assumed to arise from the interaction between a cognitive system, corresponding to the standard model in economics, and an affective system that which values earnings on a particular day relative to a daily income goal. This model can explain the entire set of results for a wide range of plausible parameter values. The key to generating the behavior is that the affective system's valuation of daily income is reference dependent: this leads to a daily "income effect" once the messenger surpasses the income goal, thus reducing the incentives to work hard for the rest of the day.

While we apply the model to our data on labor supply choices in a piece rate occupation, we believe that the insights from the model extend beyond this setting. The model may be relevant for many evnironments in which the deliberative system values progress towards a long-term goal, but the affective system's evaluation of outcomes depends on a short-term reference point. E.g., the evidence in Heath, Larrick, and Wu (1999) suggests that emotions triggered by missing or surpassing a goal extend to narrowly defined goals that exist in the context of exercising. In this case, we would argue that the deliberative system is mainly concerned with the long-term health benefits associated with the number of repetitions on a particular day, whereas the affective system leads to an emotional commitment to the short-term goal. Similarly, the affective system might evaluate daily work on a long-term project such as a dissertation relative to a reference outcome ("you have to write 3 pages a day, no matter what!"), while the deliberative system is primarily concerned with the long-term goal (graduation).

We have so far side-stepped the issue of where the reference point, i.e., the income target in our model, comes from.[20] A large literature in psychology assumes that individuals can set goals for themselves (see, e.g., Locke and Latham, 1990). Choosing a goal for oneself has an immediate and straightforward interpretation in our model. Assume that there is an initial period 0, where the only decision to be taken is the choice of the reference outcome r that applies for the rest of the time. Since this choice does not change affective payoffs in period 0, our model would predict that the deliberative self sets the reference outcome. Viewed from period 0, the deliberative system would like to maximize life-time utility. Therefore, it will choose a reference point to do "damage control" with respect to the affective system's impact on behavior in the future: since the reference point is going to affect the valuation of the affective system in future periods, choosing an appropriate reference point can minimize the affective system's impact on long-term outcomes. The model implies straightforward tradeoffs in setting an optimal reference point. For example, a higher reference point can increase overall effort. This comes at the cost of distorting within-day effort, however, and these must be traded off against the benefit of higher effort.

Such a formulation also raises interesting new questions. There might also be new testable predictions coming from a model of optimal reference points. For example, one intuition suggests that behavior should be approximately optimal with respect to long-term outcomes (e.g., the overall labor supply over a month), but not necessarily optimal on each day (e.g., displaying the income targeting level). Thus one could imagine that reference points may

---

[20] Recently, Koszegi and Rabin (2005) have proposed that the reference point is given by expectations. Our view here is not necessarily incompatible with theirs. It is, however, difficult to compare the two models, as the Koszegi and Rabin (2005) model is not easily translated into a dynamic setting.

be reset in the long run, responding to long-term changes in, e.g., the wage, but that short term changes in the wage lead to suboptimal behavior, viewed from the deliberative system's perspective. More fundamentally, extending the model to allow for an endogenous reference point raises the question how often a reference point can be reset. More research is needed on these questions.

Appendix A

This appendix details the technicalities behind the two fatigue models.

1. Fatigue with recovery (δ - fatigue)

Costs in period t are $c(e_t, k_t) = c(e_t + \delta \cdot k_t)$ where $k_t = e_{t-1} + \delta k_t$, $\delta < 1$, and $c(\cdot)$ is a strictly convex function.

Proposition 1 *Effort in the first period,* e_1, *solves* $c'(e_1) = (1 - \delta)w$. *The effort profile is u -shaped, with* $e_1 > e_t$ *for* $1 < t < m$ *and* $e_m > e_1$.

    Proof: The first order conditions are

$$\frac{\partial v}{\partial e_1} = w - c_1' - \delta c_2' - \delta^2 c_3' - \dots - \delta^{m-1} c_m' = 0$$

$$\frac{\partial v}{\partial e_2} = w - c_2' - \delta c_3' - \dots - \delta^{m-2} c_m' = 0$$

$$\dots$$

$$\frac{\partial v}{\partial e_m} = w - c_m' = 0$$

where $c_t' = c'(e_t + \delta k_t)$. Alternatively, by exploiting the recursive structure of these equations, they can be written as

$$\frac{\partial v}{\partial e_1} = w - c_1' + \delta \left(\frac{\partial v}{\partial e_2} - w \right) = 0$$

$$\frac{\partial v}{\partial e_2} = w - c_2' + \delta \left(\frac{\partial v}{\partial e_3} - w \right) = 0$$

$$\dots$$

$$\frac{\partial v}{\partial e_m} = w - c_m' = 0$$

Since $\frac{\partial v}{\partial e_t} = 0$ at the optimum, this simplifies to

$$\frac{\partial v}{\partial e_1} = (1-\delta)w - c_1' = 0 \qquad\qquad (A1)$$

$$\frac{\partial v}{\partial e_2} = (1-\delta)w - c_2' = 0$$

$$\dots$$

$$\frac{\partial v}{\partial e_m} = w - c_m' = 0$$

Because $k_1 = 0$, the first FOC in (A1) can be written $c'(e_1) = (1-\delta)w$, which proves the first claim of the proposition. Denote the solution to this FOC by e_1^*. The equations in (A1) also imply $c_t' = (1-\delta)w$ for all $t < m$. Thus, the argument of any c_t', which is k_t, must be equal to e_1^*. This implies

$$e_t + \delta k_{t-1} = e_1^*$$

Since k_{t-1} was the argument of c_{t-1}', we know it also equals e_1^*. Therefore, in any $1 < t < m$

$$e_t = (1-\delta)e_1^*.$$

This shows that the effort level drops after the first period, and remains constant thereafter, until the final period. In the final period the stock of fatigue is given by $k_m = e_{m-1} + \delta k_{m-2} = e_1^*$. Since e_m solves $w - c_m' = 0$, it must be greater than effort in the first period, e_1^*. This shows that the effort profile is u-shaped, completing the proof of the second claim in the proposition.

Proposition 2 *The response of effort to an increase in the wage is strictly positive for all work episodes up to $t < m$. In the final period, the response is also strictly positive, as long as the curvature in the marginal cost of effort is not too extreme. In particular, the following condition must hold:*

$$\frac{c_m''}{c_1''} \leq \frac{1}{\delta(1-\delta)}$$

Proof: In the proof of Proposition 1, we showed that $e_t = (1-\delta)e_1^*$ for $1 < t < m$. Thus, in order to determine the response of effort to a change in the wage, it is only necessary to know how e_1^* and e_m^* respond. Thus the individual's maximization problem can be rewritten

$$u = w(e_1 + (m-2)(1-\delta)e_1 + e_m) - (m-1)c(e_1) - c(e_m)$$

This yields two rather than m FOCs

$$\frac{\partial u}{\partial e_1} = w(1 + (m-2\_(1-\delta)) - (m-1)c_1' - \delta c_m' = 0$$

$$\frac{\partial u}{\partial e_m} = w - c_m' = 0$$

Letting $(1 + (m-2)(1-\delta)) = a$, we can write

$$c_1'(m-1) + \delta c_m' = wa$$

$$c_m' = w$$

Totally differentiating yields

$$\begin{bmatrix} (m-1)c_1'' + \delta^2 c_m'' & \delta c_m'' \\ \delta c_m'' & c_m'' \end{bmatrix} \begin{bmatrix} \dfrac{\partial e_1}{\partial w} \\ \dfrac{\partial e_m}{\partial w} \end{bmatrix} = \begin{bmatrix} a \\ 1 \end{bmatrix}$$

Denote the first matrix by H. Note that $|H| = c_m''(m-1)c_1'' > 0$.

Applying Cramer's rule, we find that $\partial e_1 / \partial w > 0$ unambiguously, which implies a strictly positive response of effort to an increase in the wage for all $t < m$, proving the first claim of the proposition. The derivative of effort in the final period with respect to the wage is

$$\frac{\partial e_m}{\partial w} = \frac{(m-1)c_1'' + \delta^2 c_m'' - \delta c_m''(1 + (m-2)(1-\delta)}{|H|}$$

$$= \frac{1}{|H|}\left[c_1''(m-1) - c_m''[-\delta^2 + \delta(1 + (m-2)(1-\delta))]\right]$$

$$= \frac{1}{|H|}\left[c_1''(m-1) - c_m''[\delta - \delta^2 + \delta(1-\delta)(m-2)]\right]$$

$$= \frac{1}{|H|}\left[c_1''(m-1) - c_m''[\delta(1-\delta)(1 + m - 2)]\right]$$

$$= \frac{1}{|H|}(m-1)[c_1'' - c_m''\delta(1-\delta)]$$

This derivative is positive if

$$\frac{c''_m}{c''_1} \le \frac{1}{\delta(1-\delta)}$$

This proves the second claim of the proposition.

Proposition 3: *In the case of fatigue with recovery, a myopic individual will increase effort in all periods in response to higher wages.*

Proof: A myopic individual ignores the impact of current effort choices on future payoffs, and thus chooses effort in each period according to

$$c'(e_1) = w$$
$$c'(e_t + \delta k_{t-1}) = w \text{ for all } t > 1\,'$$

in contrast to the FOCs in (A1). However, this implies $k_t = e_1$ for all $t > 1$, and consequently $e_t + \delta k_{t-1} = e_t + \delta e_1 = e_1$ for all $t > 1$. But then,

$$e_r = (1-\delta)e_1 \text{ for all } t > 1.$$

Since e_1 is increasing in w, so is labor supply in all other episodes.

3. Fatigue without recovery (β - fatigue)

Cost in period t are

$$c(e_t, k_t) = c(e_t + \beta \cdot k_t)$$

where $k_t = e_{t-1} + k_t$, $\beta < 1$, and $c(\cdot)$ is a strictly convex function.

Proposition 4 I*n the case of fatigue without recovery, if the marginal costs of effort are concave, the effort profile is strictly increasing over the workday.*

Proof: The first order conditions are

$$\frac{\partial v}{\partial e_1} = w - c'_1 - \beta[c'_2 + c'_3] = 0$$

$$\frac{\partial v}{\partial e_2} = w - c'_2 - \beta c'_3 = 0$$

$$\frac{\partial v}{\partial e3} = w - c'_3 = 0$$

From the final condition is follows that $c_3' = w$. Substituting into the FOC for e_2 and rearranging terms, we obtain

$$c_2' = (1 - \beta)w$$

Continuing this exercise, we get $c_1' = (1 - \beta)^2 w$. Again, it is clear that $c_t' = (1 - \beta)^{m-t} w$ for an arbitrary number of work episodes m. This sequence implies

$$c_t' = (1 - \beta)c_{t+1}'$$

Writing this out in full, we have

$$c'(e_t + \beta k_t) = (1 - \beta)c'(e_{t+1} + \beta k_{t+1})$$

$$c'(e_t + \beta k_t) = (1 - \beta)c'(e_{t+1} + \beta e_t + \beta k_t) \tag{A2}$$

for any $t < m$. To prove the claim in the proposition we proceed in two steps. First, we show that $e_t = e_{t+1}$ can never satisfy equation (1) given concave marginal costs. From the concavity of $c'(\cdot)$ it follows that

$$(1 - \beta)c'(e_{t+1} + \beta e_t + \beta k_t) < c'((1 - \beta)e_{t+1} + (1 - \beta)\beta e_t + (1 - \beta)\beta k_t)$$

Now, if $e_t = e_{t+1}$, the right hand side of the equation must satisfy the following inequality

$$c'((1 - \beta)e_{t+1} + (1 - \beta)\beta e_t + (1 - \beta)\beta k_t) < c'(e_t + \beta k_t).$$

These previous two equations imply a contradiction to (1). Second, since the right hand side of (1) is increasing in e_{t+1}, there is a similar contradiction for any $e_t > e_{t+1}$. This shows that the effort profile must be upward sloping over the day, proving the claim in the proposition.

Proposition 5 *In the case of fatigue without recovery, and a quadratic effort cost function, the response of effort to an increase in the wage is strictly positive, for all work episodes.*

Proof: Using the facts that $c_3' = w$, and $c_t' = (1 - \beta)^{m-t} w$ for $t < 3$, we solve for the optimal effort levels in each period.

$$e_1^* = (1 - \beta)^2 w$$

$$e_2^* = (1 - \beta)(1 - \beta(1 - \beta))w$$

$$e_3^* = [1 - \beta(1 - \beta)((1 - \beta) + 1 - \beta(1 - \beta))]w$$

Clearly, an increase in w translates into an increase in e_t^*, for all t. This generalizes to an arbitrary number of work episodes.

Proposition 6: *In the case of fatigue without recovery, a myopic individual will increase labor supply in all periods in response to higher wages.*

Proof: We show that a myopic individual chooses labor supply according to

$$c'(e_1) = w$$
$$e_t = (1 - \beta)^{t-1} e_1 \text{ for all } t > 1$$

Since e_1 is increasing in w, the claim follows immediately. We prove the result by induction. A myopic individual ignores the future consequences of his current choice, hence the FOCs for a myopic individual are

$$c'(e_1) = w$$
$$c'(e_t + \beta k_{t-1}) = w \text{ for all } t > 1$$

Thus, for work episode 2, we have

$$e_2 + \beta k_1 = e_2 + \beta e_1 = e_1$$
$$\Rightarrow e_2 = (1 - \beta) e_1$$

Similarly, for work episode 3, we have

$$e_3 + \beta k_2 = e_3 + \beta k_1 + \beta e_2 = e_1$$
$$\Rightarrow e_3 = e_1 - \beta k_1 - \beta e_2 = (1 - \beta) e_2 = (1 - \beta)^2 e_1$$

This provides us with the basis for the induction. Now, assume that it is true for episode m-1 that

$$e_{m-1} = (1 - \beta)^{m-2} e_1.$$

We know that $e_t + \beta k_{t-1} = e_1$. Applying this to period m, this also be written as

$$e_m + \beta k_{m-2} + \beta e_{m-1} = e_1.$$

Using the same logic for period m-1, we have

$$e_m = e_1 - \beta k_{m-2} - \beta e_{m-1}$$
$$= (1 - \beta) e_{m-1} = (1 - \beta)^{m-1} e_1.$$

This completes the proof.

Appendix B

This appendix presents the technicalities behind the example discussed in the text. The example involves a two-period version of the affect model, where the worker is below the earnings goal in the first period and above in the second.

Proposition 7: *Sufficient conditions for a maximum are given by* $[(\frac{w}{1+h})hw(v_1'' + v_2'') - c_1''] < 0$

and $c_1''(1+h) > w^2 h v_1''$.

Proof: Use the objective function, which is equation (6) in the text, and assume the optimal e_1 is in the range such that $e_2^A = \frac{r - e_1 w}{w}$, as discussed in the example. The first order conditions are then given by

$$c_1' = w\frac{1 + hv_1' + hv_2'}{1 + h} \quad \text{and} \quad c_2' = w\frac{1 + hv_2'}{1 + h}$$

The matrix of second derivatives is

$$M = \begin{bmatrix} [(\frac{w}{1+h})hw(v_1'' + v_2'') - c_1''] & \frac{w}{1+h}[hv_2''w] \\ [\frac{w}{1+h}hv_2''w] & [\frac{w}{1+h}hv_2''w - c_2''] \end{bmatrix}$$

A sufficient condition for a maximum is that the matrix be negative definite. This gives us a first condition, proving the first part of the proposition:

(B1) $\quad [(\frac{w}{1+h})hw(v_1'' + v_2'') - c_1''] < 0$

The second condition is $|M| > 0$. Writing out the determinant of M, this is given by

$$[\frac{w}{1+h}hv_2''w - c_2''][(\frac{w}{1+h})hw(v_1'' + v_2'') - c_1''] - \left[\frac{w}{1+h}(hv_2''w)\right]^2 > 0$$

$$-c_2''[(\frac{w}{1+h})hw(v_1'' + v_2'') - c_1''] + \frac{w}{1+h}(hv_2''w)[(\frac{w}{1+h})hw(v_1'' + v_2'') - c_1''] - \left[\frac{w}{1+h}(hv_2''w)\right]^2 > 0$$

Letting J denote $[(\frac{w}{1+h})hw(v_1'' + v_2'') - c_1'']$ and K denote $\frac{w}{1+h}(hv_2''w)$, rewrite as follows

$$-c_2'' J + K[J - K] > 0$$

The worker is above the target in the second period, so $v_2'' < 0$ and $K < 0$. If condition (i) from above is satisfied, then $J < 0$, and a sufficient condition for the inequality to hold is $J - K < 0$

$$[(\frac{w}{1+h})(hv_1'' w + hv_2'' w) - c_1''] - \frac{w}{1+h}(hv_2'' w) < 0$$

$$w^2 h(v_1'' + v_2'') - c_1''(1+h) < w^2 hv_2''$$

(B2) $c_1''(1+h) > w^2 hv_1''$.

This just says that MC must be steeper than MV, to have a maximum. This proves the second part of the proposition.

Proposition 8: A *sufficient condition for effort in the first period to increase, and effort in the second period to decrease, in response to a wage increase, is as follows:*

$$-1 > \frac{whv_2''(e_1 + e_2)}{1 + hv_2'} > -1 - h\frac{v_1' + wv_1'' e_1}{1 + hv_2'}$$

Proof: Totally differentiating the first order conditions yields

$$\begin{bmatrix} [(\frac{w}{1+h})(hv_1'' w + hv_2'' w) - c_1''] & \frac{w}{1+h}[hv_2'' w] \\ \frac{w}{1+h}[hv_2'' w] & [\frac{w}{1+h}[hv_2'' w] - c_2''] \end{bmatrix} \begin{bmatrix} \frac{\partial e_1}{\partial w} \\ \frac{\partial e_m}{\partial w} \end{bmatrix} = \begin{bmatrix} -[\frac{1 + hv_1' + hv_2'}{1+h}] - \frac{w}{1+h}[hv_1'' e_1 + hv_2''(e_1 + e_2)] \\ -\frac{1 + hv_2'}{1+h} - \frac{w}{1+h}hv_2''(e_1 + e_2) \end{bmatrix}$$

Simplify notation by letting q_{11} and q_{22} denote the second derivatives of the objective function with respect to e_1 and e_2, and denoting the upper and lower terms in the right hand side matrix by Z and Y. The matrix equation can then be written

$$\begin{bmatrix} q_{11} & \frac{w}{1+h}[hv_2'' w] \\ \frac{w}{1+h}[hv_2'' w] & q_{22} \end{bmatrix} \begin{bmatrix} \frac{\partial e_1}{\partial w} \\ \frac{\partial e_m}{\partial w} \end{bmatrix} = \begin{bmatrix} Z \\ Y \end{bmatrix}$$

Denoting the first matrix by K, note that the conditions from Proposition 1 imply $\mid K \mid > 0$. Applying Cramer's rule, the derivatives of first and second period effort levels with respect to the wage are given by

$$\frac{\partial e_1}{\partial w} = \frac{Det\begin{bmatrix} Z & \frac{w}{1+h}[hv_2''w] \\ Y & q_{22} \end{bmatrix}}{|K|} \quad \text{and} \quad \frac{\partial e_2}{\partial w} = \frac{Det\begin{bmatrix} q_{11} & Z \\ \frac{w}{1+h}[hv_2''w] & Y \end{bmatrix}}{|K|}$$

In order to have $\frac{\partial e_1}{\partial w} > 0$ and $\frac{\partial e_2}{\partial w} < 0$, the signs of the determinants of the two numerator matrices must be positive and negative, respectively. Writing out the determinants, these conditions can be stated

$$q_{22}Z - \frac{w}{1+h}[hv_2''w]Y > 0 \quad \text{and} \quad q_{11}Y - \frac{w}{1+h}[hv_2''w]Z < 0$$

Because $q_{11}, q_{22} < 0$, sufficient conditions for both inequalities to hold is $Z < 0$ and $Y > 0$.

$Z < 0$ is written

$$-[\frac{1 + hv_1' + hv_2'}{1+h}] - \frac{w}{1+h}[hv_1''e_1 + hv_2''(e_1 + e_2)] < 0$$

$$-[\frac{1 + hv_1' + hv_2'}{1+h}] < \frac{w}{1+h}[hv_1''e_1 + hv_2''(e_1 + e_2)]$$

(B3) $-1 - hv_1' - whv_1''e_1 < hv_2' + whv_2''(e_1 + e_2)$

$Y > 0$ is given by

$$-\frac{1 + hv_2'}{1+h} - \frac{w}{1+h}hv_2''(e_1 + e_2) > 0$$

$$-\frac{1 + hv_2'}{1+h} > \frac{w}{1+h}hv_2''(e_1 + e_2)$$

$$-1 - hv_2' > whv_2''(e_1 + e_2)$$

(B4) $-1 > \frac{whv_2''(e_1 + e_2)}{1 + hv_2'}$.

Conditions (B3) and (B4) are both satisfied if

$$-1 > \frac{whv_2''(e_1 + e_2)}{1 + hv_2'} > -1 - h\frac{v_1' + wv_1''e_1}{1 + hv_2'}$$

This proves the claim of the proposition.

References

Baumeister, Roy F. and Kathleen D. Vohs (2003). "Willpower, Choice, and Self-Control," in George Loewenstein, Daniel Read and Roy F. Baumeister, eds., Time and Decision: Economic and Psychological Perspectives on Intertemporal Choice. New York: Russell Sage Foundation, 201-216.

Beauregard, Mario, Levesque, Johanne, and Pierre Bourgouin (2001) "Neural Correlates of Conscious Self-Regulation of Emotion." *The Journal of Neuroscience*, 21, 1-6.

Benhabib, Jesand Alberto Bisin (2004). "Modelling Internal Commitment Mechanisms and Self-Control: A Neuroeconomics Approach to Consumption-Saving Decisions." Mimeo, New York University.

Bernheim, B. Douglas and Antonio Rangel (2003). "Emotions, Cognition, and Savings: Theory and Policy." Mimeo, Stanford University.

Bernheim, B. Douglas and Antonio Rangel (2004). "Addiction and Cue-Triggered Decision Processes." *American Economic Review*, 94(5), 1558-1590.

Bertrand, Marianne; Duflo, Esther and Mullainathan, Sendhil (2004). "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics*, *119*(1), 249-75.

Browning, Martin, Angus Deaton and Margaret Irish (1985) "A Profitable Approach to Labor Supply and Commodity Demands Over the Life Cycle," *Econometrica*, 53, pp. 503-43.

Camerer, Colin, Linda Babcock, George Loewenstein and Richard Thaler (1997). "Labor Supply of New York City Cabdrivers: One Day at a Time." *Quarterly Journal of Economics* 112(2), 407-41.

Chen, Keith M., Venkat Lakshminarayanan, and Laurie Santos (2005). "The Evolution of Our Preferences: Evidence form Capuchin Monkey Trading Behavior," unpublished manuscript, Yale university.

Chou, Yuan K. (2002). "Testing Alternative Models of Labor Supply: Evidence from Cab Drivers in Singapore." *The Singapore Economic Review 47(1),* pp. $17 - 47$.

Cohen, Jonathan D. (2005). "The Vulcanization of the Human Brain: A Neural Perspective on Interactions Between Cognition and Affect and Optimality in Decision Making." Working Paper, Department of Psychology, Princeton University.

De Koning, Jos, Maarten F. Bobbert and Carl Foster (1999), "Determination of optimal pacing strategy in track cycling with an energy flow model." *Journal of Science and Medicine in Sports* 2(3), $266 - 277$.

De Koning Jos, Carl Foster, Joanne Lampen, Floor Hettinga, Maarten Bobbert M (2005), "Experimental evaluation of the power balance model of speed skating." *Journal of Applied Physiology* 98, 227-233.

Farber, Henry (2005) "Is Tomorrow Another Day? The Labor Supply of New York City Cab Drivers," *Journal of Political Economy*, 113, 46-82.

Fehr, Ernst and Lorenz Goette (2006). "Do Workers work more when Wages are High? Evidence from a Randomized Field Experiment." Forthcoming, *American Economic Review*.

Foster, Carl, Jesus Hoyos, Conrad Earnest, and Alejandro Lucia (2005), "Regulation of Energy Expenditure during prolonged Athletic Competition.", *Medicine and Science in Sports and Exercise*, pp. 670 – 675.

Foster, Carl, Jos de Koning, Floor Hettinga, Joanne Lampen, Kerry La Clair, Christopher Dodge, Maarten Bobbert, and Jon Porcari (2003a), "Pattern of Energy Expenditure during Simulated Competition.", *Medicine and Science in Sports and Excercises*, pp. 826 – 831.

Foster, Carl, Jos de Koning, Floor Hettinga, Joanne Lampen, Christopher Dodge, Maarten Bobbert, and Jon Porcari (2003b), "Effect of Competitive Distance on Energy Expenditure during Simulated Competition.", *International Journal of Sports Medicine* 25, pp. 198 – 204.

Fudenberg, Drew and David Levine (2004) "A Dual Self Model of Impulse Control." Mimeo, Harvard University.

Glover, Bob and Shelly-Lynn Glover (1999) *The Competitive Runner's Handbook*, Penguin: London.

Goette, Lorenz, David Huffman and Ernst Fehr (2004) "Loss Aversion and Labor Supply," Journal of the European Economic Association, 2(2-3), 216-228.

Goette, Lorenz and Marion Lienhard (2006) "The Precautionary Effort Motive: Theory and Evidence." Manuscrip, *University of Zurich.*

Goldmark, J. and M. Hopkins (1920) "Studies in Industrial Physiology: Fatigue in Relation to Working Capacity." *Health Bulletin NO. 106.* Washington D.C.: U.S. Public Health Service.

Heath, Chip, Richard Larrick, and George Wu (1999). "Goals as Reference Points." *Cognitive Psychology* 38, pp. 79-109.

Heilizer, Fred (1977) "A Review of Theory and Research on the Assumptions of Miller's Response Competitions Model: Response Gradients," The Journal of General Psychology, 97, 17-71.

Hull, Clark (1934) "The Rat's Speed of Locomotion Gradient in the Approach to Food," *Journal of Comparative Psychology*, 17, 393-422.

Kahneman, Daniel and Amos Tversky (1979) "Prospect Theory: An Analysis of Decisions Under Risk," *Econometrica*, 47, 263-291.

Kivetz, Ran, Oleg Urminsky, and Yuhuang Zheng (2005) "The Goal Gradient Hypothesis Resurrected: Purchase Acceleration, Illusionary Goal Progress, and Customer Retention," forthcoming in the *Journal of Marketing Research*.

Kivetz, Ran, Oleg Urminsky, and Yuhuang Zheng (2005) "The Goal Gradient Hypothesis Resurrected: Purchase Acceleration, Illusionary Goal Progress, and Customer Retention," forthcoming in the *Journal of Marketing Research*.

Koszegi, Botond and Matthew Rabing (2005) "A Model of Reference-Dependent Preferences," unpublished manuscript, UC Berkeley.

Lazear, Edward (2000) "Performance Pay and Productivity." *The American Economic Review*, 90(5), 1346-1361.

Loewenstein, George, Ted O'Donoghue and Matthew Rabin (2003) "Projection Bias in Predicting Future Utility," *Quarterly Journal of Economics*, 118(4), pp. 1209 – 1248.

Loewenstein, George and Ted O'Donoghue (2005) "Animal Spirits: Affective and Deliberative Processes in Human Behavior," unpublished manuscript, Cornell University.

Locke, Edwin A. & Latham, Gary P. (1990). *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice-Hall.

MacLean, Paul D. (1990). The Triune Brain in Evolution: Role in Paleocerebral Function. New York: Plenum.

Manuck, Stephen B., Janine D. Flory, Matthew F. Muldoon, and Robert E. Ferrell (2003). "A Neurobiology of Intertemporal Choice," in George Loewenstein, Daniel Read and Roy F. Baumeister, eds., Time and Decision: Economic and Psychological Perspectives on Intertemporal Choice. New York: Russell Sage Foundation, 139-172.

Massey, Douglas S. (2002) "A Brief History of Human Society: The Origin and Role of Affect in Social Life." *American Sociological Review*, 67, (1), 1-29.

McClure, Samuel M., David Laibson, George Loewenstein, and Jonathan D. Cohen (2004). "Separate Neural Systems Value Immediate and Delayed Monetary Rewards." *Science*, 306(Oct 15), 503-507.

Metcalfe, Janet and Walter Mischel (1999). "A Hot/Cool-System Analysis of Delay of Gratification: Dynamics of Willpower." *Psychological Review* 106(1), 3-19.

Mischel, Walter, Ebbe B. Ebbesen, and Antonette Zeiss (1972). "Cognitive and Attentional Mechanisms in Delay of Gratification." *Journal of Personality and Social Pychology*, 21(2), 204-218.

Mischel, Walter, Ozlem Ayduk, and Rodolfo Mendoza-Denton (2003). "Sustaining Delay of Gratification over Time: A Hot-Cool Systems Perspective," in George Loewenstein, Daniel Read and Roy F. Baumeister, eds., Time and Decision: Economic and Psychological Perspectives on Intertemporal Choice. New York: Russell Sage Foundation, 175-200.

Mischel, Walter, Yuichi Shoda, and Monica L. Rodriguez (1989). "Delay of Gratification in Children." *Science*, 244(4907), 933-938.

Oettinger, Gerald S. (1999) "An Empirical Analysis of the Daily Labor Supply of Stadium Vendors," *Journal of Political Economy*, 107(2), 360-92.

Oi, Walter (2000) "Effort, Productivity, and Pay." mimeo, University of Rochester.

Paarsch, Harry J. and Bruce S. Shearer (2005) "The Response of Worker Effort to Piece Rates: Evidence from a Field Experiment," unpublished manuscript, Cité Universitaire, Quebec: Département d'économique, Université Laval.

Sanfey, A., Rilling, J., Aronson, J., Nystrom, D., and D. Cohen (2003) "The Neural Basis of Decision Making in the Ultimatum Game." *Science,* 300, 1254 – 1258.

See, Kelly E., Chip Heath and Craig Fox (2003) "Motivating individual performance with challenging goals: Is it better to stretch a little or a lot?" Working Paper, Fuqua School of Business.

Shearer, Bruce (2004), "Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment." *Review of Economic Studies* 71(2), 513 – 534.

Shefrin, Hersh and Richard Thaler (1988). "The Behavioral Life-Cycle Hypothesis." *Economic Inquiry,* 26, 609-643.

Shidara, Munetaka, Thomas G. Aigner, and Barry J. Richmond (1998) "Neuronal Signals in the Monkey Ventral Striatum Related to Progress through a Predictable Series of Trials," *The Journal of Neuroscience*, 18(7), 2613-2625.

Shidara, Muntetaka and Barry J. Richmond (2002) "Anterior Singulate: Single Neuronal Signals related to Degreee of Reward Expectancy," *Science*, 296, 1709-1711.

Shiv, Baba, George Loewenstein, Antoine Bechara, Hanna Damasio, and Antonio Damasio (2005). "Investment Behavior and the Dark Side of Affect." Mimeo, University of Iowa.

Shiv, Baba, Loewenstein, George, Bechara, Antoine, Damasio, Hanna, and Antonio Damasio (2005) "Investment Behavior and the Dark Side of Emotion." *Psychological Science*, 16 (June), 435-439.

Thaler, Richard H. and Hersh M. Shefrin (1981). "An Economic Theory of SelfControl." *Journal of Political Economy*, 89(2), 392-406.

Trebel, John (2003) "Intertemporal Substitution of Effort: Some Empirical Evidence." *Economica*, 70, 579-595.

Tversky, Amos and Daniel Kahneman (2000). *Choices, Values, and Frames*. Cambridge, MA: Cambridge University Press.

Vernon, Horatio (1921) *Industrial Fatigue and Efficiency*. London: George Routledge and Sons, Ltd.

Wu, George, Chip Heath, and Richard Larrick (2002) "A Value-Function Based Model of Goal Behavior," unpublished manuscript, Univ. of Chicago Graduate School of Business.

Table 1

Firm A (Swiss): Hours on Duty and Hours Active

Hours On-duty	Frequency	Percent	Hours Active	Frequency	Percent
1	64	0.34%	1	72	0.38%
2	83	0.44%	2	160	0.84%
3	460	2.43%	3	877	4.63%
4	3,949	20.84%	4	4,699	24.80%
5	9,571	50.50%	5	9,233	48.73%
6	3,262	17.21%	6	3,424	18.07%
7	1,562	8.24%	7+	483	2.55%

Firm C (Swiss): Hours on Duty and Hours Active

Hours On-duty	Frequency	Percent	Hours Active	Frequency	Percent
1	30	0.63%	1	55	1.15%
2	36	0.76%	2	103	2.16%
3	94	1.97%	3	253	5.31%
4	607	12.75%	4	858	18.02%
5	1,881	39.5%	5	1,972	41.41%
6	1,202	25.24%	6	1,299	27.28%
7	912	19.15%	7+	220	4.62%

Firm B (U.S.): Hours on Duty and Hours Active

Hours On-duty	Frequency	Percent	Hours Active	Frequency	Percent
5	0	0%	5	91	2.68%
6	22	0.65%	6	156	4.59%
7	49	1.45%	7	472	13.88%
8	154	4.55%	8	1,058	31.11%
9	689	20.34%	9	1,244	36.58%
10	1,817	53.63%	10	365	10.73%
11	639	18.86%	11	15	0.44%
12	18	0.53%	12	0	0%

Firm D (U.S.): Hours on Duty and Hours Active

Hours On-duty	Frequency	Percent	Hours Active	Frequency	Percent
5	0	0%	5	91	2.68%
6	45	1.7%	6	147	5.44%
7	96	3.62%	7	280	10.37%
8	261	9.84%	8	547	20.25%
9	676	25.48%	9	804	29.77%
10	1,088	41.01%	10	665	24.62%
11	487	18.36%	11	244	9.03%
12	0	0%	12	0	0%

Notes: Hours on duty counts all hours between a messenger's first hour with nonzero revenues and the last (includes intermediate hours with zero revenues). Active hours includes only hours with positive revenues.

Table 2a: Response of Hourly Revenues to an Increase in the Commission Rate, OLS Estimates

	Dependent Variable: Revenues in work hour k			
	Firm A hires rookie messengers, (commission rate increases at approx.12 weeks)		Firm B hires experienced messengers, (commission rate increases at approx.14 weeks)	
Δ in total daily labor supply after increase in commision rate (sum over all γ^k)	4.59** (1.019) [7.136]	1.829 (2.244) [5.25]	38.724** (4.751) [20.777]	- 8.257* (5.097) [10.662]
F-test: all γ^k are zero (p-value) 1st: clustering on date 2nd: clustering on messenger	$p <0.001$ $p <0.001$	$p <0.001$ $p <0.001$	$p <0.001$ $p <0.001$	$p <0.001$ $p <0.001$
Control Variables				
Work hour profile	Yes**,##	Yes**,##	Yes**,##	Yes**,##
Experience profile	Yes**,##	Yes**,##	Yes**	Yes**,##
Messenger fixed effect (# of messengers)	No (260)	Yes**,## (260)	No (25)	Yes**,## (25)
Hourly fixed effects (# of hours)	Yes**,## (17,964)	Yes**,## (17,964)	Yes**,## (5,499)	Yes**,## (5,499)
R^2	0.592	0.612	0.258	0.388
Number of work episodes	131,573	131,573	36,277	36,277

Notes: Commission rate increases after approximately 12 weeks at Firm A, 14 weeks at Firm B.
Table shows Δ in labor supply, measured by change in individual hourly revenues, after commission rate increases.
 a) Full regression results available from the authors upon request.
 b) In parentheses, robust s.e., adjusted for clustering on date. *, ** indicates significance at 10 and 5 percent level.
 c) In brackets, robust s.e., adjusted for clustering on messengers. #, ## indicates significance at the 10 and 5 percent level.

Table 2b: Response of Hourly Revenues to "Placebo Cutoff", OLS Estimates

	Dependent Variable: Revenues in work hour k			
	Firm C hires rookie messengers (no change in commission rate; placebo at 12 weeks)		Firm D hires experienced messengers (no change in commission rate; placebo at 14 weeks)	
Δ in daily total daily labor supply after Placebo sum over all γ^k	-4.823 (10.01) [12.34]	1.052 (13.09) [17.02]	-11.131 (9.17) [21.79]	- 15.05* (9.77) [19.8]
F-test: all γ^k are zero (p-value) 1st: clustering on date 2nd: clustering on messenger	p <0.001 p <0.001	p <0.001 p <0.001	p <0.001 p <0.001	p <0.001 p <0.001
Control Variables				
Work hour profile	Yes**,##	Yes**,##	Yes**,##	Yes**,##
Experience profile	Yes**,##	Yes**,##	Yes	Yes
Messenger fixed effect (# of messengers)	No (121)	Yes**,## (121)	No (25)	Yes (25)
Hourly fixed effects (# of hours)	Yes**,## (3,272)	Yes**,## (3,272)	Yes**,## (6,491)	Yes**,## (6,491)
R^2	0.444	0.478	0.410	0.413
Number of work episodes	32,866	32,866	28,353	28,353

Notes: Commission rate does not change at control Firms C and D.
 Table shows Δ in labor supply after "placebos": same tenure cutoffs that trigger commission rate increases at Firms A and B.
 a) Full regression results available from the authors upon request.
 b) In parentheses, robust s.e., adjusted for clustering on date. *, ** indicates significance at the 10 and 5 percent level.
 c) In brackets, robust s.e., adjusted for clustering on messengers. #, ## indicates significance at the 10 and 5 percent level.
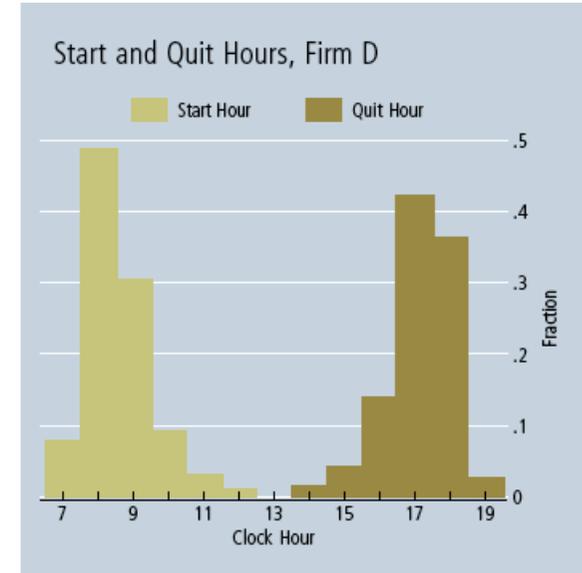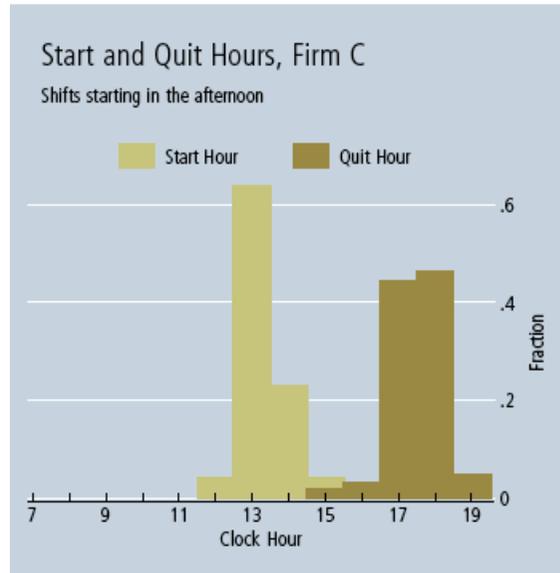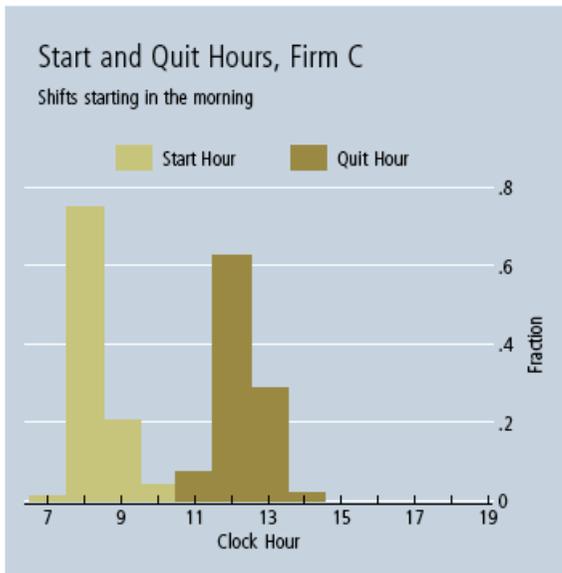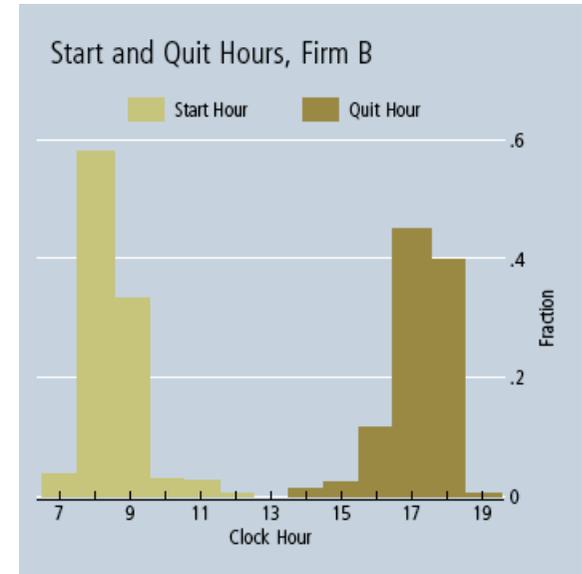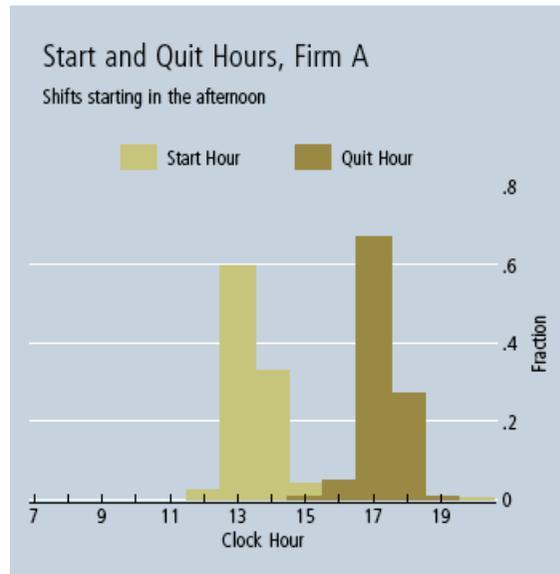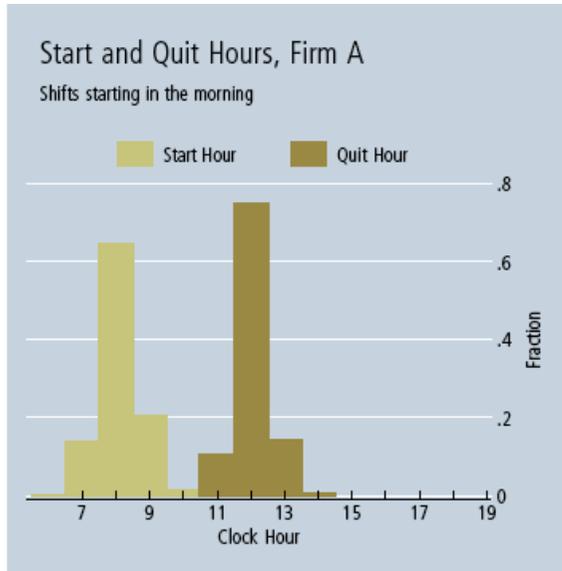
Figure 1

**How important are the following factors to be in a good
mood, rather than a neutral or bad mood, on a
particular day?**

(0 = not important at all, 1 = important, 2 = very important)

**Importance Scores (mean of responses,
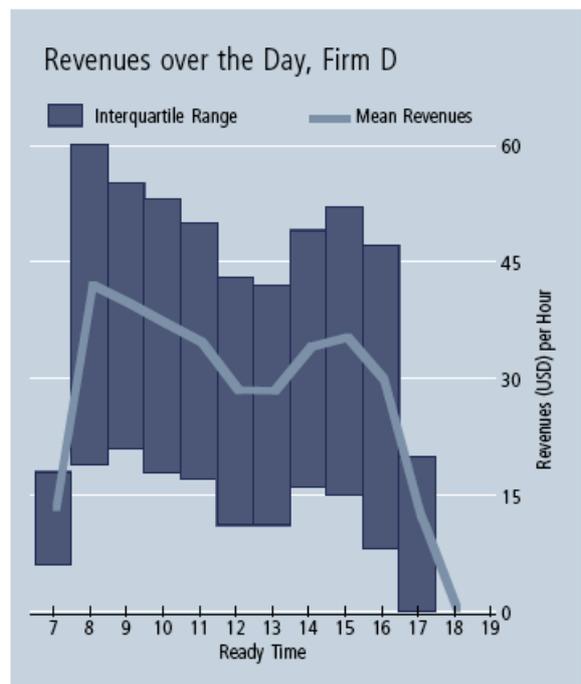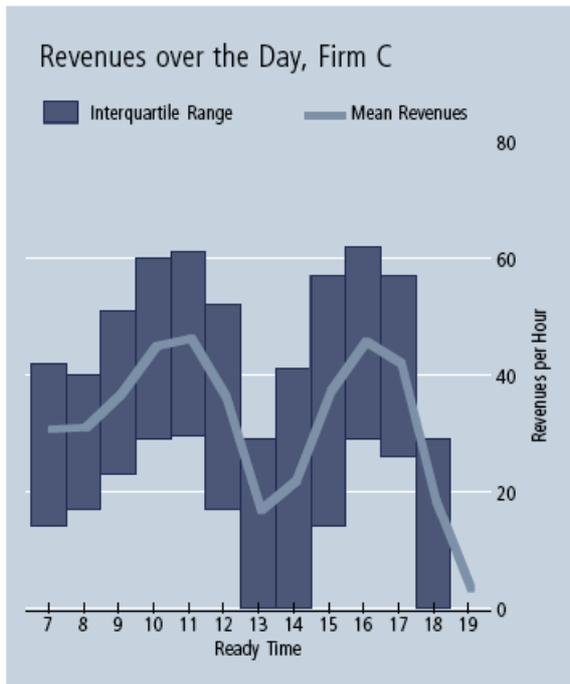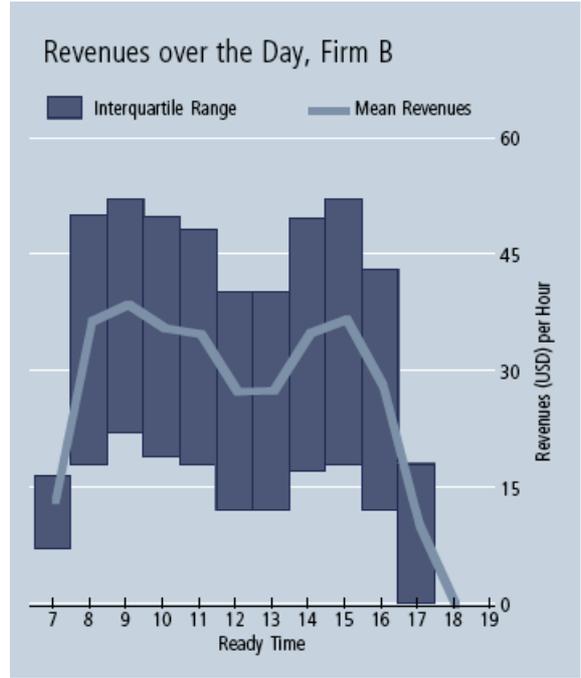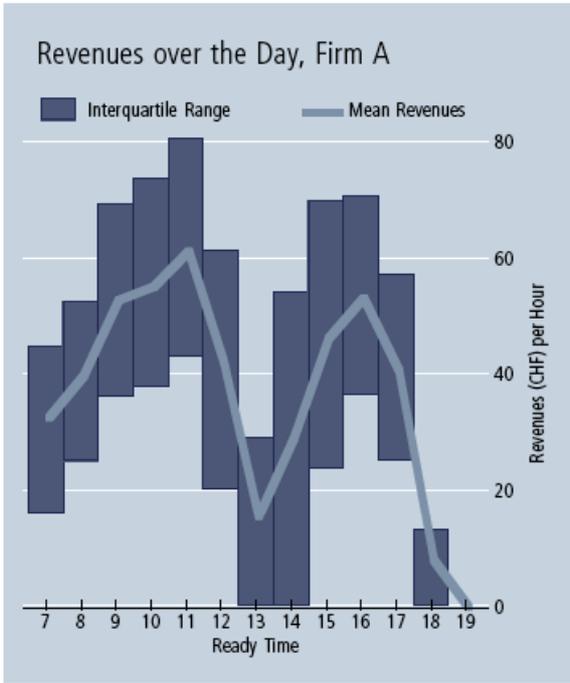+/- standard error of the mean)**
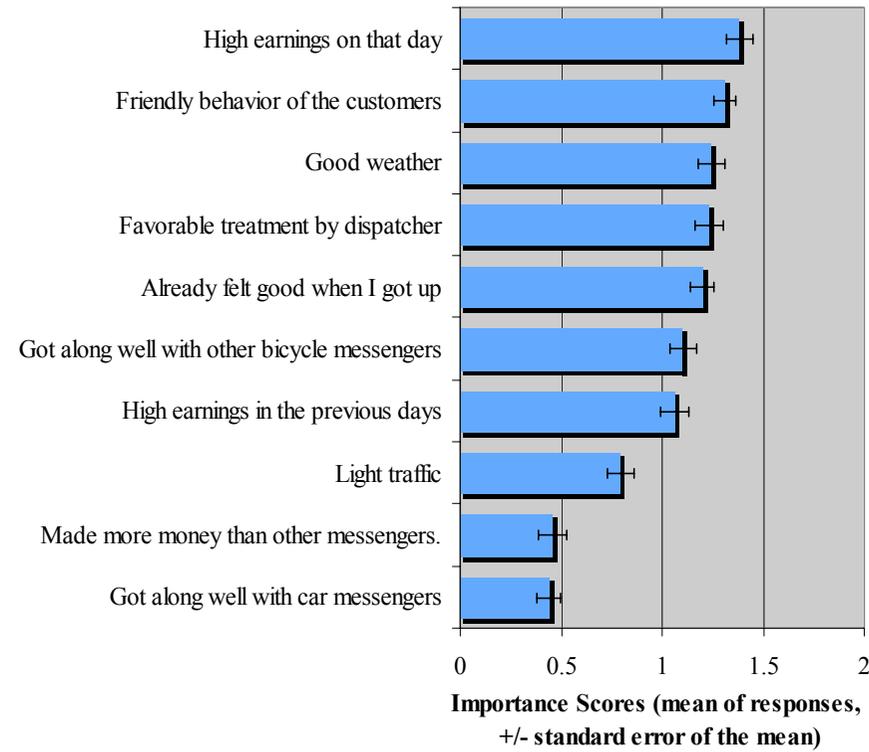
Notes: Survey of 119 bicycle messengers working in San Francisco, California.
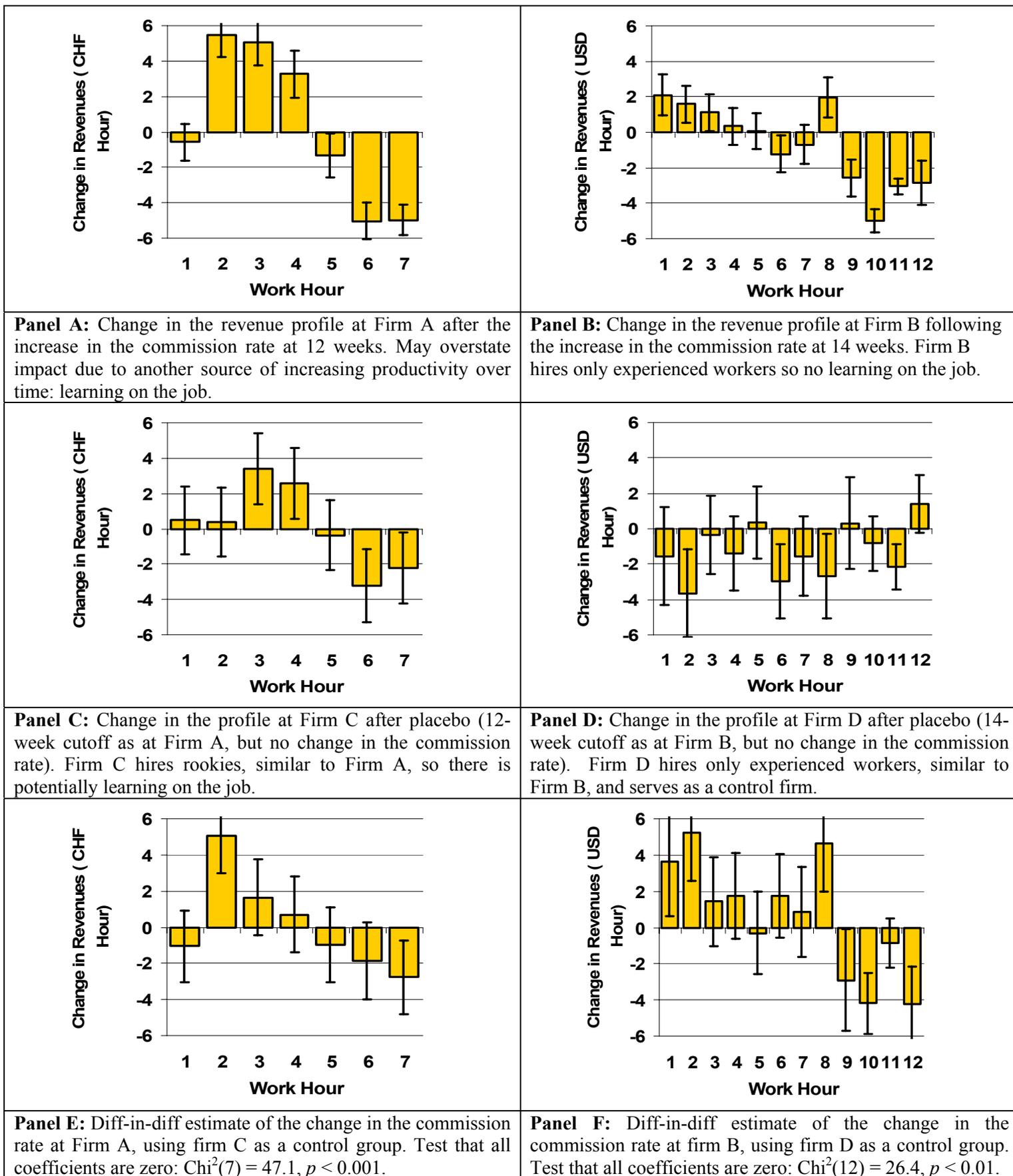
Figure 2

Figure 3

Notes: Average revenues by clock hour, normalized by average revenues in the first hour.

Figure 4: Change in Revenue Profile after Increase in Commission Rate or "Placebo Cutoff"

Panel A: Change in the revenue profile at Firm A after the increase in the commission rate at 12 weeks. May overstate impact due to another source of increasing productivity over time: learning on the job.

Panel B: Change in the revenue profile at Firm B following the increase in the commission rate at 14 weeks. Firm B hires only experienced workers so no learning on the job.

Panel C: Change in the profile at Firm C after placebo (12-week cutoff as at Firm A, but no change in the commission rate). Firm C hires rookies, similar to Firm A, so there is potentially learning on the job.

Panel D: Change in the profile at Firm D after placebo (14-week cutoff as at Firm B, but no change in the commission rate). Firm D hires only experienced workers, similar to Firm B, and serves as a control firm.

Panel E: Diff-in-diff estimate of the change in the commission rate at Firm A, using firm C as a control group. Test that all coefficients are zero: $Chi^2(7) = 47.1$, $p < 0.001$.

Panel F: Diff-in-diff estimate of the change in the commission rate at firm B, using firm D as a control group. Test that all coefficients are zero: $Chi^2(12) = 26.4$, $p < 0.01$.
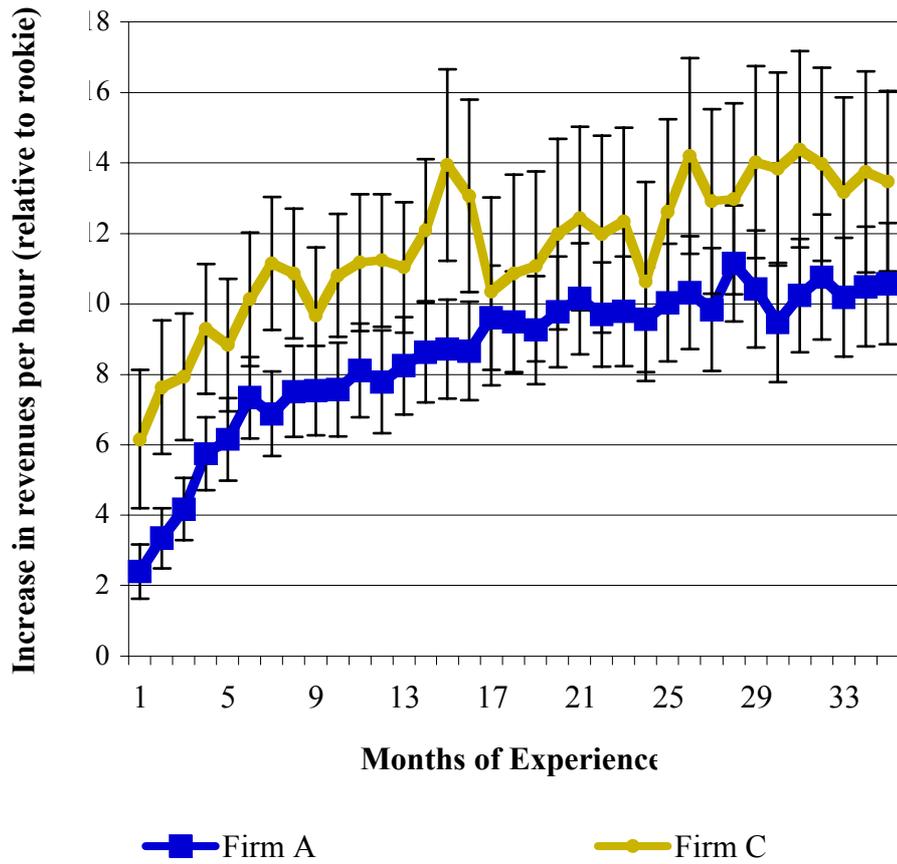
Notes: Figures in top four panels show regression coefficients from Columns (2) and (4) of Tables 2a and 2b. Error bars are standard errors of the estimate, adjusted for clustering on days.

Figure 5: Hourly Revenues Before Commission Rate Increase or "Placebo Cutoff"
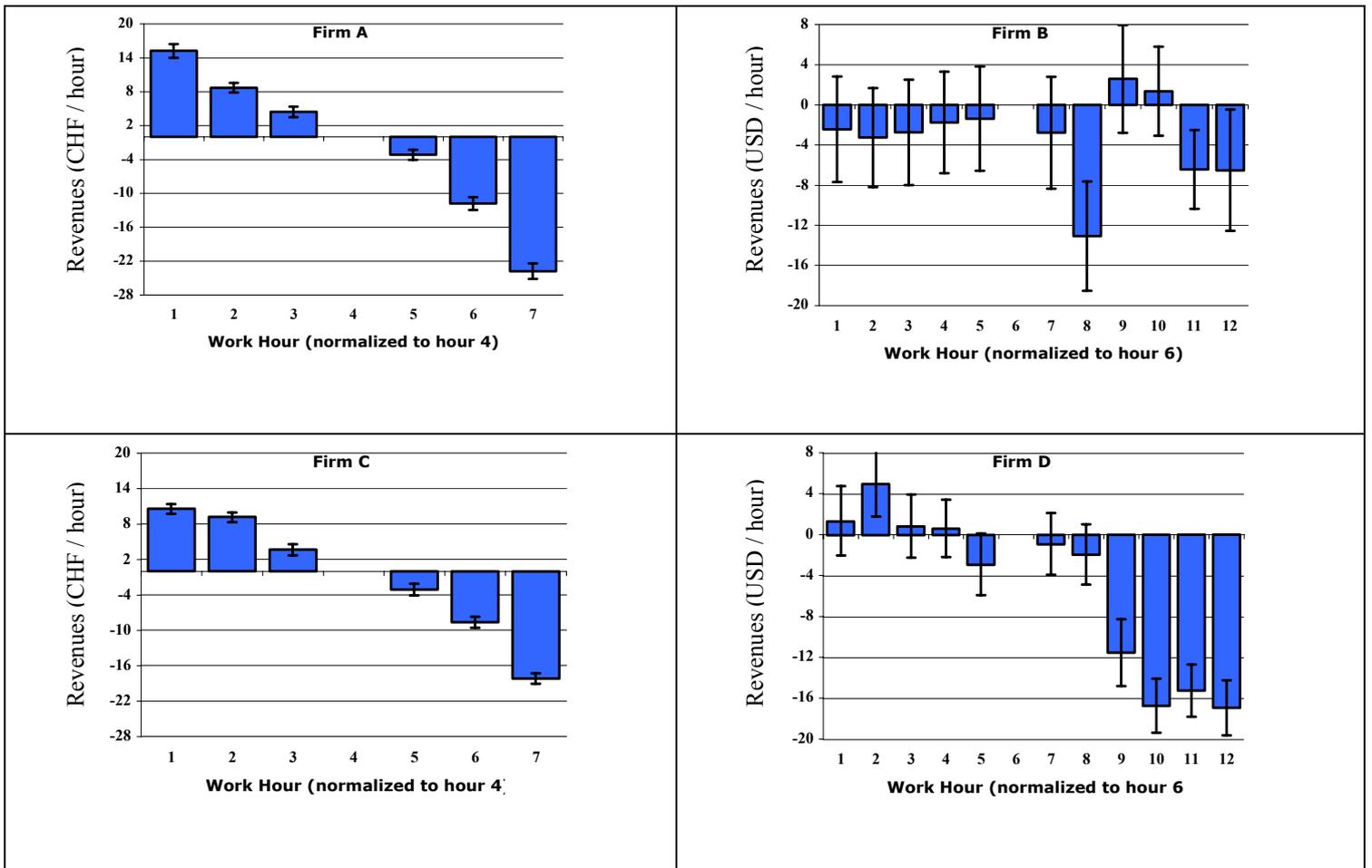
Notes: Revenues per hour before the change in piece rates at Firms A and B, and before the tenure cutoffs of 12 weeks or 14 weeks (placebos) at Firms C and D, respectively. Error bars are bars are standard errors of the estimate, adjusted for clustering on days.

Figure 6: Revenues per Hour as a Function of Experience

Notes: Regression coefficients from Column (2) of Tables 2a and 2b, respectively. Error bars are standard errors of the estimate, adjusted for clustering on dates.