

# LOSS AVERSION AND LABOR SUPPLY

---

**Lorenz Goette**  
University of Zurich

**Ernst Fehr**  
University of Zurich and CEPR

**David Huffman**  
Institute for the Study of Labor

## Abstract

In many occupations, workers' labor supply choices are constrained by institutional rules regulating labor time and effort provision. This renders explicit tests of the neoclassical theory of labor supply difficult. Here we present evidence from studies examining labor supply responses in "neoclassical environments" in which workers are free to choose when and how much to work. Despite the favorable environment, the results cast doubt on the neoclassical model. They are, however, consistent with a model of reference-dependent preferences exhibiting loss aversion and diminishing sensitivity. (JEL: J22, B49)

---

## 1. Introduction

An employer's struggle to elicit effort from employees is complicated by the fact that the timing of effort is often crucial. The employer's demand for labor effort may be extremely inelastic at certain times, when the marginal revenue product of labor is particularly high. Furthermore, because effort is typically difficult to observe, the employer must rely on incentives to elicit the desired allocation of effort over time.

Incentives are the essence of economics, and the standard model makes a clear prediction: workers work harder when there is a transitory wage increase, and substitute leisure for labor when the wage is low. This prediction has direct implications for compensation schemes, for example, employers can tie compensation to output, and thus encourage high effort at times when the firm's output is in high demand.

The first aim of this paper is to contrast the predictions of the standard model with the predictions of a model that is more firmly grounded in psychol-

---

Acknowledgments: We would like to thank George Akerlof and Matthew Rabin for helpful discussions.

E-mail addresses: Goette: [lorenz@iew.unizh.ch](mailto:lorenz@iew.unizh.ch); Huffman: [huffman@iza.org](mailto:huffman@iza.org); Fehr: [efehr@iew.unizh.ch](mailto:efehr@iew.unizh.ch)

ogy, a model of reference dependent preferences (RDP). The RDP model starts from the observation from psychology, that people tend to evaluate outcomes as gains or losses relative to a reference point, or put another way, as successes or failures relative to a goal (for a review, see Heath, Larrick, and Wu (1999); Tversky and Kahneman (2000)). The tendency to have goals is pervasive, and could naturally extend to the workplace, where workers may have personal goals in mind, in terms of income or output, when they decide how hard to work. Incorporating RDP into the standard model is simple, and leads to strikingly different predictions from the standard model in terms of how workers allocate effort over time: workers may actually work less hard on days when the wage is high.

The second aim of the paper is to assess the evidence from six recent studies of intertemporal substitution. The studies differ from the earlier literature because they focus on work environments in which workers are free to choose hours and effort; but these are environments in which the RDP model is likely to be relevant. We find that the RDP model can explain the evidence from each of the studies, without any ad hoc extensions from case to case. In particular, the RDP model can explain why higher financial incentives make workers more likely to show up for work, but at the same time can cause them to put in less effort on the job. The RDP model can also explain why an increase in the wage, or small windfall gains, distort the allocation of effort within a day, causing workers to first increase, and then decrease effort. We conclude that the weight of the evidence favors the RDP model, so that RDP should be considered an important part of understanding daily labor supply.

## 2. Theory

The standard economic model assumes that individuals are forward-looking maximizers, and that they maximize discounted lifetime utility subject to a lifetime budget constraint. Denote labor supply between  $t$  and  $t + \Delta$  by  $e_t$  and let  $e_t = 0$  if the individual is not working during period  $t$ .

Then if utility is time-separable across periods, intertemporal maximization<sup>1</sup> implies that the utility from working in period  $t$  can be represented as

$$V_t(e_t) = \lambda(w_t e_t + z_t) - g(e_t; \lambda p_t) \quad (1)$$

where  $w_t$  is the discounted wage per unit of effort  $e$  in period  $t$ ,  $\lambda$  is the marginal utility of lifetime wealth,  $z_t$  is income unrelated to effort in period  $t$ , and  $p_t$  is the

---

1. This is a simple reinterpretation of the first order conditions of lifetime utility maximization. See the Appendix in Fehr and Goette (2002) for a proof.

discounted price of consumption. The function  $g(\cdot)$  is the money-metric disutility of effort, which is increasing and convex in  $e$ .

This representation highlights the key property of any dynamic model: Small changes in income must be valued linearly. This follows from intertemporal maximization; any additional income is used to smooth out consumption over the rest of life and thus does not affect  $\lambda$ . In other words, to a first order approximation, small changes in income have no income effects. Two key implications follow directly from the linearity of period-utility in income: First, small changes in  $z_t$  do not affect labor supply. Second, temporary increases in wages must lead to higher labor supply.

### 2.1. Reference-Dependent Preferences

The psychology literature suggests that the previous model leaves out a fundamental aspect of preferences: Individuals tend to evaluate outcomes as gains and losses relative to a reference point, and thus may especially dislike a low daily income, because it feels like a loss. More formally, reference dependent preferences have two important features that influence the valuation of outcomes: *loss aversion*, which is the tendency for individuals to feel more strongly about avoiding a loss of one unit than making a gain of one unit; and *diminishing sensitivity*, which is the decrease in the marginal valuation of another unit of the outcome as the distance from the reference point increases. These two features are captured in what has become known as the Kahneman–Tversky (KT) value function.<sup>2</sup>

The Kahneman–Tversky value function is relevant for labor supply if workers have a reference income, or income goal in mind. In fact, experimental evidence shows that goals are a pervasive aspect of human decision making, and furthermore that goals, even if arbitrarily assigned, inherit the properties of the value function (Heath, Larrick, and Wu 1999). Heath, Larrick, and Wu (1999) discuss one experiment that is especially germane for labor supply: Workers seem to be twice as willing to provide an additional amount of effort to meet a given goal, than they are willing to provide the same amount of effort to surpass that goal.

The most ideal context to test the standard, neoclassical model of intertemporal labor supply is given by an environment in which workers are free to choose when and how much to work and in which there is a salient relation between their effort and their income. We hypothesize that in such environments workers are likely to have a behaviorally relevant income target. If, for example, there is a salient link between daily effort and daily income it seems likely that

---

2. The predictions of the KT value function have been borne out in many experiments and field applications. See Tversky and Kahneman (2000) for a review of the evidence.

workers have a daily income target that acts as a reference point for their daily labor supply. A parsimonious way to model behaviorally relevant income targets is to assume that individuals maximize

$$V_t(e_t) = v(w_t e_t + z_t - r) - g(e_t; \lambda p_t) \quad (2)$$

where  $v(\cdot)$  shares the properties of the KT value function around zero and  $r$  is the daily income goal that serves as the reference point. The KT value function exhibits, in particular, a discrete drop in the slope of  $v(\cdot)$  if  $w_t e_t + z_t - r$  becomes positive. If individuals value daily income according to the value function  $v(\cdot)$  in (2), instead of linearly as in (1), the marginal valuation of income, and hence the incentive to generate income through effort, changes during the course of a day. If, for example, individuals are close to, but below, the income target they face strong effort incentives whereas if they have surpassed the target the incentive to provide effort is much weaker. Thus, the first new prediction of the RDP model is that higher wages  $w_t$  may lead to lower daily effort because at high wages it is easier to surpass the income target and, therefore, the discrete drop in the marginal valuation of income occurs earlier during the day. Overall, this may lead to lower earnings during the day. The second prediction is that windfall gains  $z_t$  can affect labor supply: The individual may work harder immediately after receiving a windfall gain, because her daily income goal is suddenly within reach, but then reduce effort substantially once she surpasses her income target. In other words, windfall gains may affect the time pattern of effort during the day. The third prediction is that an increase in the wage, or a windfall gain in the morning, increases the probability of quitting early, because surpassing the income target reduces the marginal valuation of working another hour.

### 3. Evidence

#### 3.1. *The Extensive Margin*

This subsection discusses evidence from data on the extensive (participation) margin of labor supply. With regard to the participation decision the RDP model also predicts that higher temporary wages increase the temporary participation rate because at higher wages  $V_t(e_t)$  is higher. The empirical results show that workers have a positive wage elasticity of participation, that is, they are more likely to come to work on days when the wage is high.

Oettinger (1999) analyzes data on the participation decisions of vendors at a baseball stadium, during one baseball season. The vendors are independent contractors, so they are free to work or not, as they choose, at any given game. Oettinger estimates the probability of participation, conditional on the wage.

Vendors are paid on commission, so the effective wage varies across games, due to demand shocks (driven primarily by game attendance), and supply shocks (number of vendors that choose to participate). The main innovation of the study is a good set of ex-ante predictors of game attendance, which can be used to instrument for the wage: temperature, day of the week, the ranking of the home team, the quality of the opposing team, etc. In his IV estimates, Oettinger finds positive and significant wage elasticities of participation, ranging from 0.53 to 0.64.

Fehr and Goette (2002) also find a positive and significant wage elasticity of participation, in a field experiment using bicycle messengers in Zurich. Like stadium vendors, bicycle messengers are paid on commission, and they have substantial discretion over how much they work. Messengers have some fixed shifts during a month, but they can sign up for more shifts as they choose. In the experiment, there are two groups of messengers—Group A and Group B. The treatment is a month-long increase in the commission rate, of 25%. In September, Group A received the treatment and Group B was the control. In November, Group A was the control, and Group B received the treatment. Estimating the probability of working on a given day, Fehr and Goette find that messengers receiving the exogenous wage increase are significantly more likely to participate than messengers in the control group, during both months of the experiment. The implied wage elasticities, ranging from 0.72 to 0.82, are very similar to those in Oettinger (1999).

The evidence from both studies is consistent with intertemporal substitution and also with the RDP model. The next section turns to evidence on the intensive margin of labor supply, which is more helpful for distinguishing between the two models.

### 3.2. *The Intensive Margin*

*3.2.1. Effort and Across-Day Variation in the Wage.* Camerer et al. (1997) and Farber (2003) look at the response of hours worked by New York City cabdrivers to variation in the average daily wage across days, and Chou (2002) does the same for cabdrivers in Singapore. Cabdrivers are good subjects for studying intertemporal substitution on the intensive margin, because they have freedom to choose their effort during the day, and because wages vary across days due to demand shocks (conferences, rainy days, etc.).

The cabdriver studies regress log hours on the average daily wage, where the wage is calculated by dividing daily earnings by daily hours. Camerer et al. (1997) and Chou (2002) report negative wage elasticities of daily hours worked. They instrument for the wage with the average daily wage of other drivers, because measurement error in the drivers' hours could lead to a downward-biased wage elasticity. However, they find that instrumenting for the wage

actually makes elasticities even more negative, which suggests that measurement error in hours does not explain the negative elasticities. Farber is not able to instrument for the wage in his estimation, because his data include very few observations on drivers that work on the same day. All three studies find elasticities that are negative and significant. Some of the elasticities are very close to  $-1$ . The implication is that cabdrivers work long hours on days when the wage is low, and short hours on days when the wage is high.

These results are consistent with Model (2) and contradict the standard model. However, there are two main problems with the methodology of the cabdriver studies. One concern is that supply, rather than demand, is the most important source of wage variation for cab drivers. If there are common, supply-side shocks (e.g., most drivers don't want to work on the 4th of July) then the supply of cabdriver hours will be small on these days and the wage will be high. As a result, there will be a negative correlation between wage and hours that has nothing to do with RDP. A second concern is a possible selection effect. If there is a positive correlation between fixed costs of participation, and marginal disutility of effort, then on high-wage days there will be relatively more workers with a high marginal disutility of effort. This selection effect implies that effort is lower on high-wage days. Both of these concerns could lead to a downward biased estimate of the wage elasticity, but it is not clear how relevant they are empirically.<sup>3</sup>

In their field experiment, Fehr and Goette (2002) study the effort decisions of bicycle messengers. Bicycle messengers are paid on commission, so earnings are a function of effort. Messengers are relatively free to choose their effort during a day, because deliveries are called out over the airwaves to be heard by all messengers, and messengers ask for deliveries or not, as they choose. Fehr and Goette regress log revenue per shift on a dummy for the treatment—the 25% increase in the commission rate—and various controls. The controls include daily fixed effects, so the impact of the treatment is identified by comparing the treatment group to the control group, holding day characteristics constant. To control for exhaustion, the regression also includes individual-fixed effects, and controls for whether a messenger worked yesterday and whether they will work tomorrow.

Fehr and Goette find that revenues during a shift are significantly lower for the group with higher commission rates, in both months of the experiment. The implied elasticities range from  $-0.23$  to  $-0.29$ . This result is suggestive of the RDP model, in which an increase in the commission rate can lead to lower within-shift revenues, because workers reach their target, and reduce effort, earlier in a shift.

---

3. Camerer et al. (1997) and Chou (2002) argue that because cabdrivers in NYC and Singapore have regular schedules, it is unlikely that either supply shocks or selection effects can be a large factor in explaining negative elasticities.

Fehr and Goette (2002) avoid some of the methodological criticisms of the cabdriver studies with their experimental design, and address others with empirical tests. The experimental design ensures that the supply-shocks concern does not apply to the bicycle messenger results, because the source of wage variation is clear—an experimental increase in the commission rate. There could, in principle, be a selection problem, but the authors exploit the distinction between fixed and sign-up shifts to test the validity of this concern. Messengers had already chosen fixed shifts well before the announcement of the experiment, so there cannot be selection into these shifts in response to the experimental wage increase. Fehr and Goette find that revenues are lower in both fixed and sign-up shifts, with no significant difference between the two, which indicates that selection cannot explain the reduction in effort under the treatment condition.

Although the baseline results in Fehr and Goette (2002) are suggestive of the RDP model, the experimental design does not rule out another, alternative explanation. In the case of preferences that are not fully time separable, it may be optimal for workers to increase the number of shifts, but reduce effort during each shift, during a block of time in which preferences are inseparable.<sup>4</sup>

In order to distinguish between the inseparable-preferences explanation and the RDP model, the authors conduct a follow-up experiment that measures the degree of loss aversion of the individual messengers.<sup>5</sup> The strategy is to test whether the degree of loss aversion predicts the observed reduction in effort under the treatment—a distinct prediction of the RDP model. Comparing the subsample of messengers who are loss-averse to those with no indication of loss aversion, the authors find that the effort elasticity is negative and significant only for those who are loss averse. This result is shown in Figure 1(a), in which the distribution of daily revenues for the loss-averse messengers is shifted to the left, relative to the distribution of the control group. Messengers who show no indication of loss aversion, on the other hand, have an elasticity that is not significantly different from zero, which is borderline consistent with the standard model. Figure 1(b) shows that the distribution of daily revenues for messengers with no indication of loss aversion is indistinguishable from the distribution of the control group.

*3.2.2. Within-Day Labor Supply.* Farber (2003) uses new data on New York City cabdrivers that allow him to study within-day labor supply decisions.

---

4. A disadvantage of this model is that it could equally well predict a reduction in the number of shifts, and an increase in within-shift effort.

5. The measures of loss aversion are lotteries that present messengers with different choices between losses and gains, e.g., win 8 CHF with probability 0.5, or lose CHF 5 with probability 0.5. Rejection of these lotteries indicates loss aversion; as is pointed out in Rabin (2000), if final wealth is the argument in the individuals' von Neumann–Morgenstern utility function, risk aversion cannot explain an unwillingness to play for such low stakes.

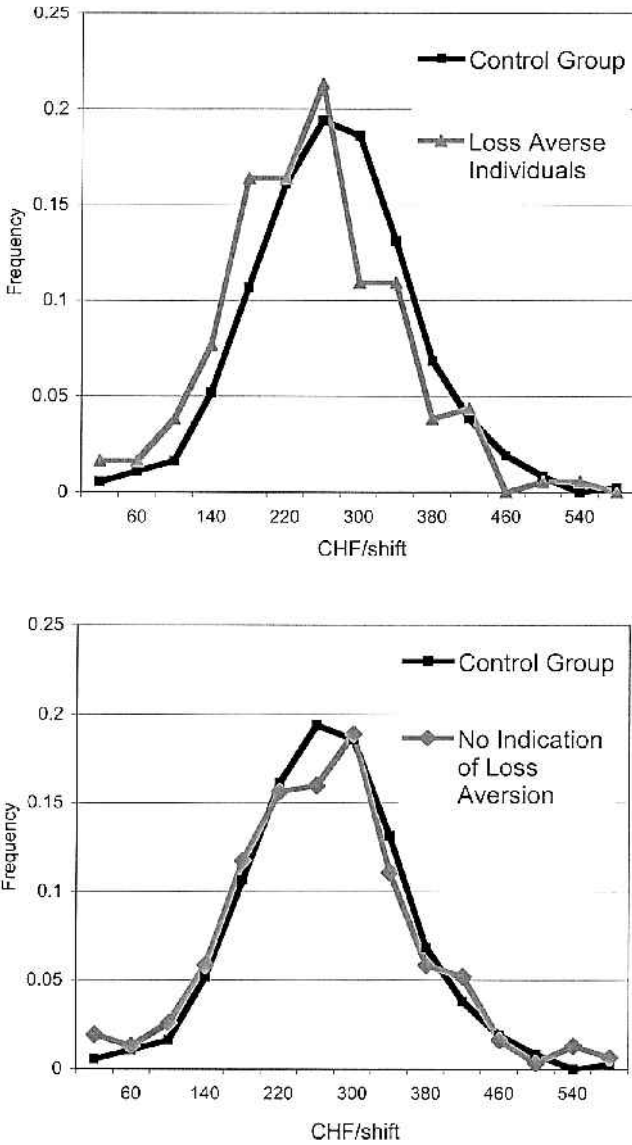


FIGURE 1. The distribution of daily revenues during the field experiment. Source: Fehr and Goette (2002)

Farber tests one of the predictions of the RDP model, that within-day earnings should affect the timing of quitting. In particular, as discussed in Section 2, high earnings early in the day should increase the probability of quitting later on in the day.



Farber uses a hazard model of quitting, and finds mixed results. His baseline results show that, conditional on hours worked, higher previous earnings increase the probability of quitting early, which is consistent with RDP. However, when Farber adds controls for driver-fixed effects, day of the week, and clock hour, he finds that previous earnings during the day no longer affect the likelihood of quitting, leaving hours worked as the main determinant of quitting. Farber also estimates the model for the subset of five drivers for whom he has at least 40 shift observations. In this case he finds that, for three out of five drivers, higher earnings significantly increase the likelihood of quitting.<sup>6</sup>

The apparent impact on Farber's results of including clock hour suggests one explanation for the mixed evidence in the paper. When Farber adds clock hours, income is no longer a significant predictor of the probability of quitting; but looking at Farber's sample, average hourly wages have a distinct profile over the day, steadily increasing over time, and peaking during the evening rush hour. They are the most predictable source of variation in hourly wages, and thus important for income as well. This means that controlling for clock hour removes an important source of information for identifying RDP.

In the absence of any clock hour constraints, the standard model predicts that drivers are unlikely to quit during rush hour, because earning opportunities are high (a rational agent would presumably postpone dinner with their family until after these key hours). On the other hand, the RDP model predicts that when earnings are particularly high, workers are more likely to quit, because they have exceeded their income target and thus have a lower marginal utility of income for the rest of the day. It is noteworthy that, in Farber's data, many cabdrivers quit just as the rush hour peaks.

There is thus an identification problem. Quite likely, the income variable picks up the effect of loss aversion and RDP when clock hour is not in the quitting regression. However, in Farber's strictest specification for the model, this effect of income on quitting is largely absorbed by the clock hour dummies. There are good reasons to control for clock hour effects: drivers may be more likely to quit during a particular hour of the day because of family obligations or because they hold second jobs. But at the same time, this removes much of the variation in earnings and quitting that could be used to identify RDP.

Goette and Huffman (2003) study the within-day labor supply decisions of bicycle messengers. Their data are the delivery records of three bicycle messenger firms, two of which are located in San Francisco, California and a third that is located in Basel, Switzerland. Messengers at these firms are paid on commission, and have substantial discretion over choice of effort.

Goette and Huffman test two fairly subtle predictions of the RDP model:

---

6. This is consistent with the lottery results from Fehr and Goette (2002), and survey evidence from Goette and Huffman (2003), which suggest that not all workers are loss averse.

(1) A windfall gain in the morning should lead to increasing, and then decreasing effort over the course of the afternoon, compared to a worker without a windfall gain; (2) a worker on a high wage should exhibit higher effort, and then lower effort over the course of the day, compared to a worker on a low wage. Both of these predictions follow from the fact that individuals with RDP have a higher incentive to provide effort when they are close to, but below, their target whereas when they have reached the target the effort incentives are lower.

To test the first prediction, Goette and Huffman study the impact of earnings during the morning on effort at different points in time in the afternoon. Their regression also includes messenger-fixed effects, firm-day-clock hour-fixed effects, start hour, and other controls, so that all individual-specific, firm-day-and-hour-specific and start-time-specific variation in morning earnings is removed.<sup>7</sup> The authors argue that the remaining variation in morning earnings is unrelated to effort and mainly captures luck—being at the right place at the right time.<sup>8</sup> Thus, the interaction terms capture the impact of windfall gains in the morning on effort in each hour of the afternoon. Figure 2 is a graphical representation of these regression results, showing the estimated change in revenues during each hour of the afternoon caused by a \$50 windfall gain in the morning. At both firms, the observed pattern is exactly as predicted by the RDP model: A windfall gain in the morning leads to significantly higher effort early in the afternoon, but significantly lower effort in the early evening, relative to a messenger without windfall gains.

Goette and Huffman test the second prediction using Firms B and C, which raise the commission rate for messengers after they have been working for a few months. They study how a higher commission rate affects effort during each hour of the workday.<sup>9</sup> The regressions include the usual fixed effects, start hour, and other controls. Figure 3 presents the results of these regressions, showing the impact of an increase in the commission rate from 50% to 55% at Firm B, and from 38% to 43% at firm C. At both firms (one in California and the other in Switzerland), messengers on the higher commission rate work significantly

---

7. The remaining variation is significant: the standard deviation of the residual from a regression of morning earnings on fixed effects represents 34% of average morning earnings.

8. For at least two reasons, luck is very important for earnings. First, earnings vary with the service type and the geographic pickup and dropoff zones of the delivery—so that earnings are substantially influenced by being in the right place at the right time. For example, being near the border of Zone 2 when a rush delivery is called out going from a nearby location in Zone 1, to Zone 2, would result in earnings that are roughly 60% higher than average hourly earnings. Another source of randomness, often cited by messengers, comes from good fortune in getting deliveries that line up. Making a delivery to a distant location can be very profitable, if other deliveries come up during the ride that can be taken care of without deviating too far from the “line,” or route, between the messenger’s origin and final destination.

9. At Firm C, messengers work shorter shifts, which may be in the morning or afternoon. So, for comparison purposes, the authors focus on hour of work rather than clock hour.

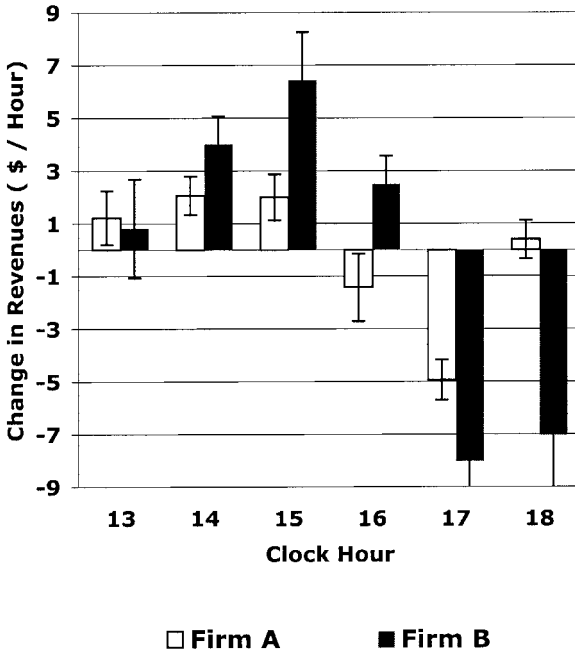


FIGURE 2. The impact of a \$50 increase in morning earnings on messengers' hourly revenues in the afternoon ( $\pm$ s.e. of estimate). Source: Goette and Huffman (2003)

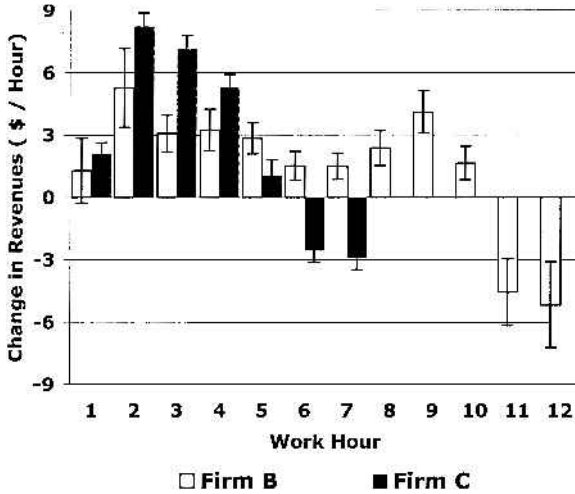


FIGURE 3. The impact of a 5-percentage-point increase in the commission rate on messengers' hourly revenues during the workday ( $\pm$ s.e. of estimate). Source: Goette and Huffman (2003)

harder during early work hours, but significantly less hard during later work hours, as predicted by the RDP model.

Goette and Huffman note that, although messengers on a higher commission rate work at a slower pace at the end of the day, they work harder on average. This contrasts with Fehr and Goette (2002), in which a wage increase had a negative impact on within-shift productivity. However, these results need not be inconsistent. The increase in the commission rate was much larger in Fehr and Goette (2002), 25% rather than 10%, and the RDP model predicts the following: a larger increase in the wage leads to a correspondingly larger decrease in the time needed to reach the daily income target, and hence increases the likelihood of an overall reduction in effort.

Goette and Huffman also discuss whether fatigue could be an alternative explanation for their results. They argue that there are several reasons why fatigue is very unlikely to explain the observed differences in behavior. First, and perhaps most importantly, in a survey of 114 bicycle messengers Goette and Huffman find that the majority of messengers say they have an income target. They ask: “*After earning \_\_\_ dollars during the day, it feels less urgent to earn another dollar (if this question does not apply to you, answer with N.A.).*” Seventy-three percent respond that they have such a dollar amount in mind. Second, the data show that messengers in San Francisco have very few deliveries during the lunch hour. The authors argue that this substantial resting period in the middle of the day makes it unlikely that even a very busy morning has an impact on effort costs in the afternoon.<sup>10</sup> Third, the authors note that an intuitively appealing cost of effort function, in which costs are increasing in previous effort expenditures, predicts the opposite of the observed pattern. Because effort at time  $t$  increases costs at time  $t + 1$ , this cost function implies that a utility-maximizing messenger who has had an exhausting morning should rest early in the afternoon, and work hard at the end of the day when there is no concern about the impact of effort on future costs. Similarly, the cost function implies that a messenger with a higher commission rate should display increasing effort over the day, which is the opposite of the observed pattern.<sup>11</sup> Goette and Huffman conclude that the pattern of increasing and then decreasing effort over the day reflects RDP rather than fatigue.

---

10. One reason why the analysis divides the day into morning (7:00–12:00) and afternoon (13:00–18:00) is because, at all firms, there is a dramatic drop in deliveries during the lunch hour.

11. An unusual cost function in which effort at time  $t$  decreases effort costs at  $t + 1$  could explain why messengers work hard early in the afternoon, and less hard at the end of the day. But this function cannot explain Goette and Huffman’s second result: the cost function implies that messengers on a high commission rate should display decreasing effort over the day, but cannot explain why effort should drop below the effort of a messenger on the lower commission rate.

## 4. Conclusions

This paper examines the evidence on a central prediction of the standard model: When wages are temporarily high, individuals should supply more labor. Furthermore, absent significant nonseparabilities in preferences over labor supply, a wage increase should cause labor supply to increase along all dimensions.

The evidence we review establishes several results: The participation margin of labor supply indeed responds to higher wages in the predicted fashion. The evidence on effort, or within-day labor supply, is less supportive of the standard model. Several studies report negative elasticities of effort with respect to across-day variation in daily wages. The allocation of effort over the day, following a wage increase or after a small windfall gain, also appears inconsistent with the standard model.

There are two broad classes of models to explain these results: Models that rely on nonseparabilities in labor supply, and models that incorporate reference-dependent preferences. Models that rely on nonseparabilities require different assumptions to explain each set of results, and sometimes, the assumptions needed to explain one set of results generate the wrong prediction for another. In contrast, a model that relies on reference-dependent preferences is well grounded in psychology, and can explain the entire set of results. By Occam's Razor, we favor this model.

## References

- 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, 407–441.
- Chou, Yuan K. (2002). "Testing Alternative Models of Labor Supply: Evidence from Cab Drivers in Singapore." *The Singapore Economic Review*, 47, 17–47.
- Farber, Henry S. (2003). "Is Tomorrow Another Day? The Labor Supply of New York Cab Drivers." NBER Working Paper 9706.
- Fehr, Ernst and Lorenz Goette (2002). "Do Workers Work More when Wages are High? Evidence from a Randomized Field Experiment." IEW Working Paper 144.
- Goette, Lorenz and David Huffman (2003). "Reference-Dependent Preferences and the Allocation of Effort over Time: Evidence from Natural Experiments." Working paper, Institute for Empirical Research in Economics, University of Zurich.
- Heath, Chip, Richard Larrick, and George Wu (1999). "Goals as Reference Points." *Cognitive Psychology*, 38, 79–109.
- Oettinger, Gerald S. (1999). "An Empirical Analysis of the Daily Labor Supply of Stadium Vendors." *Journal of Political Economy*, 107, 360–392.
- Tversky, Amos and Daniel Kahneman (2000). *Choices, Values, and Frames*. Cambridge University Press.