LABOR SUPPLY OF NEW YORK CITY CAB DRIVERS: ONE DAY AT A TIME*

Colin Camerer, Caltech
Linda Babcock, CMU
George Loewenstein, CMU
Richard Thaler, U of Chicago

*This paper is a revised and shortened version of a paper with the same title, published in the Quarterly Journal of Economics, May 1997, 407-441. It is published in Choices, Values, and Frames, D.Kahneman and A. Tversky (Eds.), 2000, Cambridge: Cambridge University Press, pp. 356-370.

I. INTRODUCTION

Theories of labor supply predict how the number of hours people work will change when their hourly wages or income change. The standard economic prediction is that a temporary increase in wages should cause people to work longer hours. This prediction is based on the assumption that workers substitute labor and leisure intertemporally, working more when wages are high and consuming more leisure when its price-- the foregone wage-- is low (e.g., Lucas and Rapping, 1969). This straightforward prediction has proven difficult to verify. Studies of many types often find little evidence of intertemporal substitution (e.g., Laisney, Pohlmeier, and Staat, 1992). However, the studies are ambiguous because when wages change, the changes are usually not clearly temporary (as the theory requires). The studies also test intertemporal substitution jointly along with auxiliary assumptions about persistence of wage shocks, formation of wage expectations, separability of utility in different time periods, etc.

An ideal test of labor supply responses to temporary wage increases requires a setting in which wages are relatively constant within a day but uncorrelated across days, and hours vary every day. In such a situation, all dynamic optimization models predict a positive relationship between wages and hours (e.g., MaCurdy [1981] p. 1074).

Such data are available for at least one group of workers-- New York City cab drivers. Drivers face wages which fluctuate on a daily basis due to demand shocks caused by weather, subway breakdowns, day-of-the-week effects, holidays, conventions, etc. Although rates per mile are set by law, on busy days drivers spend less time searching for customers and thus earn a higher hourly wage. These wages tend to be correlated within days and uncorrelated across days (i.e., transitory).

Another advantage of studying cab drivers is that, unlike most workers, they choose the number of hours they work each day because drivers lease their cabs from a fleet for a fixed fee (or own them) and can drive as long as they like during a continuous 12-hour shift. Furthermore, most analyses of labor supply measure hours (and sometimes income) by self-reports. For cab drivers, better measures of hours and income are available from "trip sheets" the drivers fill out and from meters installed in cabs, which automatically record the fares.

Because drivers face wages which fluctuate from day to day, and can work flexible hours, the intertemporal substitution hypothesis makes a clear prediction: Drivers will work longer hours on high-wage days. Behavioral economics suggests an alternative prediction (which is what motivated our research in the first place): Many drivers told us they set a target for the amount of money they wanted to earn that day, and quit when they reached the target. (The target might be a certain amount beyond the lease fee, or twice the fee.) Daily targeting makes exactly the opposite prediction of the intertemporal substitution hypothesis: When wages are high, drivers will reach their target more quickly and quit early; on low-wage days they will drive longer hours to reach the target. To test the standard intertemporal substitution hypothesis against the daily targeting alternative, we collected three samples of data on the hours and wages of drivers.

We find little evidence for positive intertemporal substitution because most of the wage elasticities-- the ratio of percentage change in hours to percentage change in wages-- are estimated to be negative. This means that drivers tend to quit earlier on high wage days drive longer on low wage days. Elasticities for inexperienced drivers are around -1 for two of the three samples of cab drivers we used in our study. The results are robust to outliers and many different specifications. (And since our paper was originally published, in 1997, one replication using survey data from Singapore also found negative elasticities; see Chou, 2000.) There are several possible explanations for these negative elasticities, other than the daily targeting hypothesis, but most can be comfortably ruled out.

II. EMPIRICAL ANALYSES

In this section, we use data on trip sheets of New York City cab drivers to explore the relationship between hours that drivers choose to work each day and the average daily wage. Many details are omitted here but are included in Camerer et al (1997).

A trip sheet is a sequential list of trips that a driver took on a given day. For each trip, the driver lists the time the fare was picked up and dropped off and the amount of the fare (excluding tip). Fares are set by the Taxi and Limousine Commission (TLC). For the first period we study (1988), the fares were \$1.15 per trip plus \$.15 for each 1/5 of a mile or 60 seconds of waiting time. For the second period we study (1990 and 1994) fares were \$1.50 per trip plus \$.25 each 1/5 of a mile or 75 seconds of waiting time. In both periods, a \$.50 per-trip surcharge is added between 8 PM and 6 AM.

Our data consist of three samples of trip sheets. We describe each data set briefly. The first data set, TRIP, came from a set of 192 trip sheets from the spring of 1994. We borrowed and copied these from a fleet company. Fleet companies are organizations that own many cabs (each affixed with a medallion which is required to operate it legally). They rent these cabs for 12 hour shifts to drivers who, in our sample period, typically paid $76 for a day shift and $86 for a night shift. The driver also has to fill the cab up with gas at the end of the shift (costing about $15). Drivers get most of their fares by "cruising" and looking for passengers. (Unlike many cities, trips to the airport are relatively rare--around one trip per day on average). Drivers keep all the fares including tips. The driver is free to keep the cab out as long as he wants, up to the 12 hour limit. Drivers who return the cab late are fined. When a driver returns the cab the trip sheet is stamped with the number of trips that have been recorded on the cab's meter. This can then be used to determine how carefully the driver has filled in the trip sheet.

The measure of hours worked is obtained directly from the trip sheet. It is the difference between the time that the first passenger is picked up and the time that the last passenger is dropped off. Total revenue was calculated by adding up the fares listed on the trip sheet. The average hourly wage is total revenue divided by hours worked.

Many of the trip sheets were incomplete, since the number of trips listed by the cab driver was much fewer than the number of trips recorded by the meter. Therefore, we exclude trip sheets that listed a number of trips that deviates by more than two from the metered number. This screen leaves us with 70 trip sheets from 13 drivers (eight of whom drive on more than one day in the sample).

The advantage of the TRIP data set is that we can use the trip sheets to measure the within-day autocorrelation in hourly earnings as well as differences in earning across days. Even though taxi fares are fixed by the TLC, earnings differ from day to day because of differences in how "busy" drivers are -- that is, whether they spend most of the day with passengers in their cab, or have to spend a lot of time searching for passengers.

The second and third data sets of trip sheets were obtained from the TLC. The TLC periodically samples trip sheets to satisfy various demands for information about drivers and earnings (e.g., when rate increases are proposed). In these two data sets, hours and the number of driver-listed trips are obtained from the trip sheets and number of recorded trips, fares, and miles driven are obtained from the meter.

The TLC developed a screen to discard incomplete trip sheets. Because the TLC provided us with the summary measures, but not the trip sheets themselves, we are unable to create an alternative screening procedure, so we use their screened data for our analyses.

The first of the TLC data sets, TLC1, is a summary of 1723 trip sheets from 1990. This data set includes three types of drivers: Daily fleet drivers, lease drivers who lease their cabs by the week or month, and others who own a medallion-bearing cab and drive it. Most owner-drivers rent their cab out to other drivers for some shifts, imposing constraints on when and how long they can drive. Those who do not rent out their cabs can drive whenever they want.

The screened data contain 1044 trip sheets and 484 drivers (234 of whom drove more than one day in the data). The main advantages of this sample are that it includes several observations for each of many drivers and contains a range of different types of drivers.

The second TLC data set, TLC2, is a summary of 750 trip sheets, mostly from November 1-3, 1988. This data set samples owner-drivers as well as drivers from mini-fleet companies (mini-fleets usually lease cabs to drivers weekly or monthly). We discard 38 trip sheets using the TLC screen, leaving us 712 trip sheets. The main differences between TLC2 and TLC1 are that no drivers appear more than once in the data in TLC2 and the fares in TLC2 are slightly lower.

The analyses reported in the body of the paper use only the screened samples of trip sheets for all three data sets. Including the screened-out data does not make much difference.

To learn about important institutional details we also conducted a phone survey of 14 owners and managers at fleet companies which rent cabs to drivers. The average fleet in New York operates 88 cabs so the responses roughly summarize the behavior of over a thousand drivers. The survey responses help make sense of the results derived from analysis of hours and wages.

Sample Characteristics

Table I presents means, medians, and standard deviations of the key variables. Cab drivers work about 9.5 hours per day, take between 28 and 30 trips, and collect almost $17 per hour in revenues (excluding tips). In the TRIP data, the average trip duration was 9.5 minutes and the average fare was $5.13. Average hourly wage is slightly lower in the TLC2 sample because of the lower rates imposed by the TLC during that time period.

In the empirical analyses below, we estimate labor supply functions using the daily number of hours as the dependent variable and the average wage the driver earned during that day as the independent variable (both in logarithmic form). The average wage is calculated by dividing daily total revenue by daily hours. This, however, assumes that the decisions drivers make regarding when to stop driving depend on the average wage during the day, rather than fluctuations of the wage rate during the day.

Fluctuations within- and across-days are important because testing for substitution requires that wages be different and roughly uncorrelated across days (and they were), and that hourly wages be correlated within a day. We used the trip-by-trip data available in the TRIP sample to construct hour-by-hour measures of wages. One hour's median wage had an autocorrelation of .493 with the previous hour's wage, so there is indeed a strong positive correlation within each day; when a day starts out as a high wage day, it will probably continue to be a high wage day. The fleet managers surveyed weakly agreed[1] with these patterns, saying the within-day autocorrelation is positive or zero (none said it was negative). Since wages are different each day, fairly stable within days, but uncorrelated across days, they are ideal for calculating the labor supply response to a temporary changes in wages.

Wage Elasticities

The simple correlations between log hours and log wages are all modestly negative, -.503, -.391, and -.269. The wage elasticity-- the percentage change in hours relative to the percentage change in wage-- can be estimated by simply regressing the logarithm of hours against the logarithm of a worker's wage, using ordinary least squares. These regressions yield estimates between -.19 and -.62, which are generally highly significantly different from zero.

1 Fleet managers were asked whether "a driver who made more money than average in the first half of a shift" was likely to have a second half which was better than average (3 agreed), worse than average (0) or about the same as average (6). Expressing the target-income hypothesis, two fleet managers spontaneously said the second half earning were irrelevant "because drivers will quit early".

However, this standard technique can be misleading because of a potential bias caused by measurement error. Measurement error is a pervasive concern in studies of labor supply, particularly because most data are self-reports of income and hours which may be subject to memory or recording errors, or self-presentation biases. Though the data on hours come from trip sheets rather than from memory, they may still include recording errors. Unfortunately, even if errors in the measurement of hours are random, they lead to a predictable bias in the wage elasticity: Because the average hourly wage is derived by dividing daily revenue by reported hours, overstated hours will produce hours that are too high <u>and</u> wages that are too low. Understated hours will produce hours that are too low and wages that are too high. Measurement error in hours can therefore create spuriously negative elasticities. This bias can be eliminated if we can find a proxy for the drivers' wage which is highly correlated with the wage, but uncorrelated with a particular driver's measurement error in hours. (Such a proxy is called an "instrumental variable" (IV) in econometrics.) Fortunately, an excellent proxy for a driver's wage is a measure of the wage of <u>other</u> drivers who are working on the same day during the same shift.[2] We use these measures of other-driver wages in all the regressions that follow.

Regressions of (log) hours on (log) wages are shown in Table II for the three data sets. TRIP and TLC1 include multiple observations for each driver, so either the standard errors are corrected to account for the panel nature of the data, or driver fixed effects are included. A driver fixed effect is a dummy variable for each driver which adjusts for the possibility that each driver might systematically drive more or less hours, holding the wage constant, than other drivers. Several other variables controlling for weather conditions and shift dummy variables were also included; their effects were

---

[2] In fact, we used three summary statistics of the distribution of hourly wages of other drivers that drove on the same day and shift (the 25th, 50th, and 75th percentiles) as instruments for a driver's wage.

modest and are not shown in Table II.

The IV elasticities in Table II are negative and significantly different from zero, except in the TRIP sample when fixed effects are included. Indeed, the elasticities in the TLC samples are close to -1, which is the number predicted by daily targeting theory. The results in Table II are quite robust with respect to various specifications we tried to control for outliers, such as median regression. The difference between the wage elasticities in the two TLC samples and the fixed-effects estimate in the TRIP sample can be explained by a difference in the composition of types of drivers across the three samples.[3]

How do Elasticities Vary with Experience?

Drivers may learn over time that driving more on high wage days and less on low wage days provides more income and more leisure. If so, the wage elasticities of experienced drivers should be more positive than for inexperienced drivers. There are good measures of driver experience in these data sets. In the TLC data sets, the TLC separated drivers into experience groups: for TLC1, those with greater or less than four years of experience and in TLC2, those with greater or less than three years of experience. These group measures are absent in the TRIP data. However, cab driver licenses are issued with six-digit numbers (called hack numbers), in chronological order, so that lower numbers correspond to drivers who obtained their licenses earlier. Using their license numbers, we use a median split to

[3] TRIP consists entirely of fleet drivers (who pay daily) while the TLC samples also includes weekly and monthly lease-drivers, and owner-drivers. Lease-drivers and owner-drivers have more flexibility in the number of hours they drive (since fleet drivers are constrained to drive no more than 12 hours). Elasticities for the fleet drivers are substantially smaller in magnitude (less negative) than for lease- and owner-drivers (as we see below). The different results in the TRIP sample, which is all fleet drivers, reflects this compositional difference in driver types.

divide drivers into low- and high-experience subsamples for the TRIP data.

Table III presents the wage elasticities estimated separately for low- and high-experience drivers. All regressions include fixed effects (except, of course for TLC2). In all three samples, the low-experience elasticity is significantly negative, and insignificantly different from -1. The wage elasticity of the high-experience group is significantly larger in magnitude for the TRIP and TLC2 samples (p=.030 and .058 respectively), and insignificantly smaller in the TLC1 sample.

How do Elasticities Vary with Payment Structure?

The way drivers pay for their cabs might affect their responsiveness of hours to wages if, for example, the payment structure affects the horizon over which they plan. Alternatively, it might affect the degree to which they can significantly vary hours across days. The TLC1 sample contains data from three types of payment schemes -- daily rental (fleet cabs), weekly or monthly rental (lease cabs), or owned. Table IV presents elasticity estimates in the three payment categories from the TLC1 sample. All regressions are estimated using instrumental variables and include driver-fixed effects.

All wage elasticities in Table IV are negative. The elasticity which is smallest in magnitude, for fleet drivers, is not significantly different from zero. The lease and owner-driver wage elasticities are approximately -.9 and are significantly different from zero. Part of the explanation for the lower elasticity for fleet drivers is a technical one. Since they are constrained to drive no more than 12 hours, the dependent variable is truncated, biasing the slope coefficient towards zero.

Could Drivers Earn More by Driving Differently?

One can simulate how income would change if drivers changed their driving behavior. Using the TLC1 data, we take the 234 drivers who had two or more days of data in our sample. For a specific driver i, call the hours and hourly wages on a specific day t, h_{it} and W_{it}. respectively, and call driver i's mean hours over all the days in the sample $h_{i.}$ By construction, the driver's actual total wages earned in our sample is $\Sigma_t h_{it}W_{it}$.

One comparison is to ask how much money that driver would have earned if he had driven h_i hours every day rather than varying the number of hours. Call this answer "fixed-hours earnings" (FHE), $\Sigma_t h_i W_{it}$.

Is FHE greater than actual earnings? We know that, on average, h_{it} and w_{it} are negatively correlated so that the difference between FHE and actual earnings will be positive in general. In fact, drivers would increase their net earnings by 5.0 percent on average (std.error =.4 percent) if they drove the same number of hours (h_i) every day, rather than varying their hours every day. If we exclude drivers who would earn less by driving fixed hours (because their wage elasticity is positive), the improvement in earnings would average 7.8 percent. And note that if leisure utility is concave, fixed-hours driving will improve overall leisure utility too.

These increases in income arise from following the simplest possible advice -- drive a constant number of hours each day. Suppose instead that we hold each driver's average hours fixed, but reallocated hours across days as if the wage elasticity was +1. Then the average increase in net income across all drivers is 10 percent. Across drivers who gain, the average increase is 15.6 percent.

III. Explaining Negative Wage Elasticities

Wage elasticities estimated with instrumental variables are significantly negative in two out of three samples. Elasticities are also significantly higher for experienced drivers in two of three samples, and significantly more negative for lease- and owner-drivers than for fleet drivers. These two empirical regularities, along with other patterns in the data, and information gleaned from our telephone survey of fleet managers, allow us to evaluate four alternative explanations for the observed negative elasticities. Ruling out these alternatives is important (see Camerer et al, 1997 for details), because it leaves daily targeting as the most plausible explanation for anomalous negative elasticities.

One hypothesis is that drivers are "liquidity-constrained"-- they don't have much cash to pay everyday expenses (and cannot borrow), so they cannot quit early on low-wage days. But drivers who own their cab medallions are presumably not liquidity-constrained (because medallions are worth $130,000), and their elasticities are negative too.

A second possibility is that drivers finish late on low-wage days, but take lots of unrecorded breaks on those days, so they actually work fewer hours. But we excluded long breaks from the TRIP sample and found no difference in the results.

A third possibility is that drivers quit early on high-wage days because carrying a lot of passengers is especially tiring. But the fleet managers we surveyed said the opposite; most of them thought that fruitlessly searching for fares on a low-wage days was more tiring than carrying passengers.

A fourth alternative is more subtle: We only have observations of work hours on the days that drivers chose to to work at all (or "participate", in labor economics jargon). Omitting non-working days can bias the measured elasticity negatively if the tendency for a driver to work unexpectedly on a certain day is correlated with the tendency to work unusually long hours (Heckman, 1979). But drivers usually participate on a fixed schedule of shifts each week (and often must pay their lease fee, or some penalty, if they do not show up for scheduled work), so there is little unexpected participation and probably very little bias.

A fifth alternative is that drivers like happy endings: They drive until they earn a lot in a final unit of time (such as their final trip, or final hour). Ross and Simonson (1996) report evidence that people like "happy endings" and will end event sequences happily when they can. Drivers who create happy endings will drive longer on slow days (if the earnings that constitute a happy ending are not too responsive to earnings earlier in the day) than drivers on good days. We tested this hypothesis by comparing earnings in the final hour with earlier earnings, but found no evidence of a happy-ending effect.

Daily Income Targeting

As explained in the introduction, the prediction we sought to test in our study is based on two assumptions: Cab drivers take a one-day horizon, and set a target (or target range) and quit when the target is reached.

Taking a one-day horizon is an example of narrow "bracketing" (Read and Loewenstein, 1996) simplifying decisions by isolating them from the stream of decisions they are embedded in. For example, people are risk averse to single plays of small gambles, even though they typically face many uncorrelated small risks over time which diversify away the risk of a single play. Bettors at horse tracks seem to record the betting activity for each day in a separate "mental account" (Thaler, this volume). Since the track takes a percentage of each bet, most bettors are behind by the end of the day. Studies show that they tend to shift bets toward longshots in the last race in an attempt to `break even' on that day (McGlothlin, 1956). Read and Loewenstein [1995] observed an unusual kind of bracketing among trick-or-treaters on Halloween. Children told to take any two pieces of candy at a single house always chose two different candies. Those who chose one candy at each of two adjacent houses (from the same set of options) typically chose the same candy at each house. Normatively, the children should diversify the portfolio of candy in their bag, but in fact they only diversify the candy from a single house. Isolation of decisions has also been observed in strategic situations: Camerer et al (1993) found that subjects in a three-stage `shrinking-pie' bargaining experiment often did not bother to look ahead and find out how much the `pie' they bargained over would shrink if their first-stage offers were rejected.

The notion that drivers are averse to falling below a target income is consistent with other evidence that judgments and decisions depend on a comparison of potential outcomes against some aspiration level or reference point [Helson, 1949; Kahneman and Tversky, 1979; Tversky and Kahneman, 1991], and people are dispropoportionally sensitive to losing, or falling short of a reference point.[4]

Both narrow bracketing and loss-aversion are analytically necessary to explain negative wage elasticities. A one-day horizon is necessary because drivers who take a longer horizon, even two days, can intertemporally substitute between the two days and will have positive wage elasticities. Therefore, if their elasticities are negative they must be taking a one-day horizon. Aversion to falling short of the target is a necessary ingredient because if drivers do take a one-day

[4] Other applications of loss-aversion include Kahneman, Knetsch and Thaler (1990) on "endowment effects" in consumer choice and contingent valuation of nonmarket goods, Samuelson and Zeckhauser (1988) on "status quo biases", and Bowman et al (1997) and Shea (1995) on anomalies in savings-consumption patterns.

horizon, elasticities will only be highly negative if the marginal utility of daily income drops sharply around the level of average daily income, which is just a labor-supply way of saying they really dislike falling short of a daily average (compared to how much they like exceeding it).

Furthermore, the daily targeting hypothesis rang true to many of the fleet managers we surveyed. They were asked to choose which one of three sentences "best describes how many hours cab drivers drive each day?". Six fleet managers chose "Drive until they make a certain amount of money". Five chose the response "Fixed hours". Only one chose the intertemporal substitution response "drive a lot when doing well; quit early on a bad day".

Several other studies with field data have used the same ingredients-- narrow bracketing and loss-aversion-- to explain anomalies in stock market behavior and consumer purchases. For example, the “equity premium puzzle” is the tendency for stocks (or “equity”) to offer much higher rates of returns than bonds over almost any moderately long time interval, which cannot be reconciled with standard models of rational asset pricing. Benartzi and Thaler (1995) argue that the large premium in equity returns compensates stockholders for the risk of suffering a loss over a short horizon. They show that if investors evaluate the returns on their portfolios once a year (taking a narrow horizon), and have a piecewise-linear utility function which is twice as steep for losses as for gains, then investors will be roughly indifferent between stocks and bonds, which justifies the large difference in expected returns. If investors took a longer horizon, or cared less about losses, they would demand a smaller equity premium. Two experimental papers have demonstrated the same effect (Thaler, Tversky, Kahneman and Schwartz, 1997; Gneezy and Potters, 1997).

Experimental and field studies show that investors who own stocks that have lost value hold them longer than they hold `winning’ stocks, before selling (Shefrin and Statman, 1985; Odean, 1996; Weber and Camerer, in press). Purchase of consumer goods like orange juice fall a lot when prices are increased, compared to how much purchases rise when prices are cut (Hardie, Johnson and Fader, 1993). These tendencies can only be explained by investors and consumers isolating single decisions about stocks and products from the more general decision about the contents of their stock portfolio or shopping cart, and being unusually sensitive to losing money on the isolated stock or paying more for the isolated product.

Various psychological processes could cause drivers to use daily income targeting. For example, targeting is a simple decision rule: It requires drivers to keep track only of the income they have earned. This is computationally easier than tracking the ongoing balance of foregone leisure utility and marginal income utility (which depends on expected future wages) which is required for optimal intertemporal substitution. Targeting might just be a heuristic shortcut which makes deciding when to quit easier.

Daily targets can also help mitigate self-control problems (as many mental accounts do, see Shefrin and Thaler, 1992). There are two kinds of self-control problems drivers might face. First, driving a cab is tedious and tiring and, unlike many jobs, work hours are not rigidly set; drivers are free to quit any time they want. A daily income goal, like an author imposing a daily goal of written pages, establishes an output-based guideline of when to quit. A weekly or monthly target would leave open the temptation to quit early today and make up for today's shortfall tomorrow, or next week, and so on, in an endless cycle.

Second, in order to substitute intertemporally, drivers must save the windfall of cash they earn from driving long hours on a high-wage day so they can afford to quit early on low-wage days. But a drive home through Manhattan with $200-$300 in cash from a good day is an obstacle course of temptations for many drivers, creating a self-control problem that is avoided by daily targeting.

Finally, daily targeting can account for the effect of experience rather naturally: Experienced drivers who have larger elasticities either learn over time to take a longer horizon (and to resist the temptations of quitting early and squandering cash from good days), or to adopt the simple rule of driving a fixed number of hours each day. Alternatively, some drivers may just lack these qualities to begin with and they quit at higher rates, selecting themselves out of the experienced-driver pool because they have less leisure and income. Either way, experienced drivers will have more positive wage elasticities.

IV. DISCUSSION AND CONCLUSIONS

Dynamic theories of labor supply predict a positive labor supply response to temporary fluctuations in wages. Previous studies have not been able to measure this elasticity precisely, and the measured sign is often negative, contradicting the theory. These analyses, however, have been plagued by a wide variety of estimation problems.

Most estimation problems are avoided by estimating wage elasticities for taxi drivers. Drivers have flexible self-determined work hours and face wages that are highly correlated within days, but only weakly correlated between days, (so fluctuations are transitory). The fact that our analyses yield negative wage elasticities suggests that elasticities of intertemporal substitution around zero (or at least, not strongly positive) may represent a real behavioral regularity. Further support for this assertion comes from analyses of labor supply of farmers (Berg, 1961; Orde-Brown, 1946) and self-employed proprietors (Wales, 1973) who, like cab drivers, set their own hours and often have negative measured wage elasticities. These data suggest that it may be worthwhile to search for negative wage elasticities in other jobs in which workers pay a fixed fee to work, earn variable wages and set their own work hours-- such as fishing, some kinds of sales, and panhandling.

Of course, cab drivers, farmers, and small-business proprietors are not representative of the working population. Besides some demographic differences, all three groups have self-selected onto occupations with low variable wages, long hours and (in the case of farmers and cab drivers), relatively high rates of accidents and fatalities. However, there is no reason to think their planning horizons are uniquely short. Indeed, many cab drivers are recent immigrants who, by immigrating, are effectively making long-term investments in economic and educational opportunity for themselves and their children.

Because evidence of negative labor supply responses to transitory wage changes is so much at odds with conventional economic wisdom, these results should be considered a provocation for further theorizing. It may be that the cab drivers' situation is special. Or it may be that people generally take a short horizon and set income targets, but adjust these targets flexibly in ways which can create positive responses to wage increases,[5] so that myopic adjustable targeting can explain both positive elasticities observed in some studies and the negative elasticities observed in drivers.

We have two ideas for further research. A natural way to model a driver's decision is by using a hazard model which specifies the probability that a driver will quit after driving t hours, as a function of different variables observable at t. Daily targeting predicts that quitting will depend on

[5] For example, suppose the target is adjusted depending on the daily wage (e.g., a driver realizes this will be a good day and raises his target for that day). Then his behavior will be very much like that of a rational driver intertemporally substituting over time, even though the psychological basis for it is different (and does not require any foresight).

the total wages cumulated at t in a strongly nonlinear way (when the daily total reaches a target the probability of quitting rises sharply). Intertemporal substitution predicts that quitting will depend only on the average wage earned up to time t.

Another prediction derived from daily targeting is that drivers who receive an unusually big tip will go home early. Experimenters posing as passengers could actually hand out big tips (say, $50) to some drivers and measure, unobstrusively, whether those drivers quit early compared to a suitable control group. Standard theory predicts that a single large tip produces a tiny wealth effect which should not make any difference to current behavior[6], so a perceptible effect of a big tip would be more evidence in favor of daily targeting and against intertemporal substitution.

Final comments

As part of a broader project in behavioral economics, work like ours strives to draw discipline and inspiration for economic theorizing from other social sciences, particularly psychology, while respecting the twin aesthetic criteria that characterize post-war economics: models should be formal and make field-testable predictions. The goal is to demonstrate that economic models with better roots in psychology can create interesting challenges for formal modelling, and make better predictions.

---

[6] A crucial assumption is that the tip is seen by the driver as a temporary wage increase, rather than an indicator that more large tips may come in the hours ahead (which would cause them to drive longer). Controlling for drivers' beliefs, and observing their hours, are challenges for experimental design.

The ingredients of our project suggest a recipe for doing convincing behavioral economics "in the wild". We derived a simple hypothesis from behavioral economics-- daily targeting-- which predicts that the <u>sign</u> of a regression coefficient would be the opposite of the sign predicted by standard theory, so we have a dramatic difference in two theories. We got lucky and found good data. We had an excellent proxy variable (or instrument) for a driver's daily wage, the wage of other drivers working at the same time, which eliminated the bias caused by measuring hours with error. We also obtained variables which enabled us to rule out some alternative explanations (such as liquidity constraint and effects of breaks). And we found an effect of experience which is consistent with the hypothesis that targeting is a costly heuristic which drivers move away from with experience, in the direction of intertemporal substitution. Critics who think our findings of negative elastiticities are an econometric fluke must explain why we did <u>not</u> find negative elasticities for experience drivers.

Finally, a growing number of economists have begun to question the benefits of increasing sophistication in mathematical models. In game theory, theorists and experimenters have shown that simple evolutionary and adaptive models of behavior can often explain behavior better than sophisticated equilibrium concepts (e.g., John Gale, Kenneth Binmore, and Larry Samuelson, 1995; Camerer, Ho and Chong, 2001). Experimental economists have noted how "zero intelligence" programmed agents can approximate the surprising allocative efficiency of human subjects in double auctions (Dan Gode and Shyam Sunder, 1993), and how demand and choice behavior of animals duplicates patterns seen in empirical studies of humans (John Kagel, Raymond Battalio, and Leonard Green, 1995). Our research, too, shows that relatively simple principles and models can often go a long way toward explaining and predicting economic behavior, and even outperform more sophisticated models of economic agents.

References

Altonji, Joseph G., "Intertemporal Substitution in Labor Supply: Evidence From Micro Data," *Journal of Political Economy*, XCIV (1986), s176-s215.

Benartzi, Shlomo and Richard Thaler, "Myopic Loss Aversion and the Equity Premium Puzzle," *Quarterly Journal of Economics*, CX (1) (1995), 73-92.

Berg, Elliot J., "Backward-Sloping Labor Supply Functions in Dual Economies-- The Africa Case," *Quarterly Journal of Economics*, LXXV (3) (1961), 468-492.

Bowman, David, Debby Minehart, and Matthew Rabin. "Loss Aversion in a Savings Model," *Journal of Economic Behavior and Organization*, 38, February 1999, 155-178.

Browning, Martin, Angus Deaton, and Margaret Irish, "A Profitable Approach to Labor Supply and Commodity Demands Over the Life-Cycle," *Econometrica*, LIII (1995), 503-43.

Camerer, Colin F., Babcock, Linda; Loewenstein, George; and Thaler, Richard. "Labor supply of New York City cabdrivers: One day at a time," *Quarterly Journal of Economics*, May 1997, 407-441.

Camerer, Colin F. , Teck-Hua Ho, and Juin Kuan Chong, "Behavioral game theory: Thinking, learning and teaching", Caltech working paper. http://www.hss.caltech.edu/~camerer/camerer.html

Chou, Yuan K. Testing Alternative Models of Labor Supply: Evidence from Taxi Drivers in Singapore, 2000. University of Melbourne Department of Economics, http://www.economics.unimelb.edu.au/research/768.htm

Colin F. Camerer, Eric Johnson, Talia Rymon, and Sankar Sen. `Cognition and sequential bargaining over gains and losses,' in K. Binmore, A. Kirman, and P. Tani (eds.), *Frontiers of Game Theory*, MIT Press, 1993.

Duesenberry, J. *Income, Saving, and the Theory of Consumer Behavior*, (Cambridge, MA: Harvard University Press, 1949).

Gneezy, Uri and Potters, Jan, "An Experiment on Risk Taking and Evaluation Periods," *Quarterly Journal of Economics*, 112, May 1997, 631-646.

Hardie, Bruce G. S., Eric J. Johnson and Peter S. Fader, "Modeling Loss Aversion and Reference-Dependence Effects on Brand Choice," *Marketing Science*, 12, 1993, 378-394.

Heckman, James, "Sample Selection Bias as a Specification Error."*Econometrica*, XLVII (1979), 153-161.

Helson, Harry. *Adapation-Level Theory*. New York: Harper and Row, 1964.

Kahneman, Daniel, Jack Knetsch, and Richard Thaler, "Experimental Tests of the Endowment Effect and the Coase Theorem," *Journal of Political Economy*, XCVIII (1990), 1325-1348.

Kahneman, Daniel, and Amos Tversky, "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, XLVII (1979), 263-91.

Laisney, Francois, Winfried Pohlmeier, and Matthias Staat, `Estimation of Labor Supply Functions using Panel Data: A Survey', in Matyas and Sevestre, eds., *The Economics of Panel Data: Handbook of Theory and Applications* (Kluwer, 1992).

Lucas, Robert E. Jr., and Leonard A. Rapping, "Real Wages, Employment, and Inflation," *Journal of Political Economy*, LXXVII (1969).

MaCurdy, Thomas E., "An Empirical Model of Labor Supply in a Life-Cycle Setting," *Journal of Political Economy*, LXXXIX (1981), 1059-85.

Maddala, G. S. Introduction to Econometrics, 2nd edition, 1992. New York: Macmillan Publishing Company.

Mankiw, N. Gregory, Julio J. Rotemberg, and Lawrence H. Summers, "Intertemporal Substitution in Macroeconomics," *Quarterly Journal of Economics*, C (1986), 225-251.

McGlothlin, William H. `Stability of Choices Among Uncertain Alternatives,' *American Journal of Psychology*, LXIX (1956), 604-615.

Mulligan, Casey. `The Intertemporal Substitution of Work-- What Does the Evidence Say?' University of Chicago Population Research Center working paper, June 1995.

NYC Taxi and Limousine Commission. "Taxi Trip and Fare Data: A Compendium", October 29, 1991. NYC Taxi and Limousine Commission. "The New York City Taxicab Fact Book," May 1992.

Odean, Terry. `Are Investors Reluctant to Realize Their Losses?' *Journal of Financ*e, 53, October 1998, 1775-1798.

Orde-Brown, G., *Labour Conditions in East Africa*, (London: Colonial Office, H.M.S.O., 1946)

Pencavel John. "Labor Supply of Men: A Survey," in O. Ashenfelter and R. Layard (Eds.), *Handbook of Labor Economics, Volume I*. (Amsterdam, The Netherlands: North-Holland, 1986), pp. 3-102.

Pindyck, Robert S. and Daniel L. Rubinfeld. *Microeconomics* (New York: Macmillan, 1989).

Read, D. and Loewenstein, G., `The diversification bias: Explaining the difference between prospective and real-time taste for variety'. *Journal of Experimental Psychology: Applied*, 1, (1995), 34-49.

Read, D. and Loewenstein, G., "Temporal Bracketing of Choice: Discrepancies Between Simultaneous and Sequential Choice." Working Paper, 1996. Department of Social and Decision Sciences Working Paper,Carnegie Mellon University.

Ross, William T., Jr., and Itamar Simonson, " Evaluations of Pairs of Experiences: A Preference for Happy Endings," *Journal of Behavioral Decision Making, IV (* 1991), 155-61.

Samuelson, William, and Richard Zeckhauser, "Status Quo Bias in Decision Making," *Journal of Risk and Uncertainty*, I (1988), 39-60.

Shea, John, "Union Contracts and the Life-Cycle/Permanent-Income Hypothesis," *American Economic Review,* LXXXV (1) (1995), 186-200.

Shefrin, Hersh M., and Richard H. Thaler, "Mental Accounting, Saving, and Self-Control," in G. Loewenstein and J. Elster, eds., *Choice Over Time* (New York: Russell Sage Foundation Press, 1992).

Thaler, Richard. `Mental Accounting and Consumer Choice,' *Marketing Science*, IV (1985). 199-214.

Thaler, Richard, Amos Tversky, Daniel Kahneman, and Alan Schwartz. `How Myopic Loss Averse Investors Learn From Experience,' *Quarterly Journal of Economics* (in press).

Tversky, Amos, and Daniel Kahneman, "Loss Aversion in Riskless Choice: A Reference-Dependent Model," *Quarterly Journal of Economics*, CVI (1991), 1039-61.

Wales, Terence J. "Estimation of a Labor Supply Curve for Self-Employed Business Proprietors," *lnternational Economic Review*, XIV(l) (1973), 69-80.

Weber, Martin and Colin F. Camerer. `The Disposition Effect in Securities Trading: An Experimental Analysis', *Journal of Economic Behavior and Organization*, 33, January 1998, 167-184.

Table I. Summary Statistics

	Mean	Median	Std. Dev
TRIP (n = 70)			
Hours Worked	9.16	9.38	1.39
Average Wage	16.91	16.20	3.21
Total Revenue	152.70	154.00	24.99
# Trips Counted by Meter	30.70	30.00	5.72
TLC1 (n = 1044)			
Hours Worked	9.62	9.67	2.88
Average Wage	16.64	16.31	4.36
Total Revenue	154.58	154.00	45.83
# Trips Counted by Meter	27.88	29.00	9.15
TLC2 (n = 712)			
Hours Worked	9.38	9.25	2.96
Average Wage	14.70	14.71	3.20
Total Revenue	133.38	137.23	40.74
# Trips Counted by Meter	28.62	29.00	9.41

Table II. Instrumental Variable (IV) Regression of Log Hours Against Log Hourly Wage

Sample	TRIP		TLC1		TLC2
Log Hourly Wage	-.319 (.298)	.005 (.273)	-1.313 (.236)	-.926 (.259)	-.975 (.478)
Fixed Effects	No	Yes	No	Yes	No
Sample Size	70	65	1044	794	712
Number of Drivers	13	8	484	234	712

Note: Dependent variable is the log of hours worked. Other independent variables (not shown) are high temperature, rain, and dummy variables for during-the-week shift, night shift, and day shift. Standard errors are in parentheses. Instruments for the log hourly wage include the summary statistics of the distribution of hourly (log) wages of other drivers on the same day and shift (the 25th, 50th, and 75th percentiles).

Table III. IV Log Hours Regression by Driver Experience Level

Sample	TRIP		TLC1		TLC2	
Experience Level	Low	High	Low	High	Low	High
Log Hourly Wage	-.841 (.290)	.613 (.357)	-.559 (.406)	-1.243 (.333)	-1.308 (.738)	2.220 (1.942)
Fixed Effects	Yes	Yes	Yes	Yes	No	No
Sample Size	26	39	319	458	320 375	
P-value for Difference in Wage Elasticity	.030		.666		.058	

Note: See note to Table II.

Table IV. IV Log Hours Regressions by Payment Structure (TLC1 data)

Type of Cab	Fleet	Lease	Owned
Log Hourly Wage	-.197 (.252)	-.978 (.365)	-.867 (.487)
Fixed Effects	Yes	Yes	Yes
Sample Size	150	339	305

Note: See note to Table II. Fleet cabs are rented daily, leased cabs are rented by the week or month, and owned cabs are owned by the drivers.