Reference-Dependent Preferences and Labor Supply: The Case of New York City Taxi Drivers

By Henry S. Farber*

Increased attention has been paid in recent years to deviations from the standard neoclassical model of consumer behavior. A substantial segment of this work focuses on reference-dependent preferences where there is a change in the shape of the utility function at some base (reference) level of income or consumption. These models have strong predictions for how responses to changes in prices are affected by the actual level of consumption or income relative to the reference level. A difficulty in bringing this class of models to the data is that the reference level of income or consumption is seldom observed, and we have only weak information on what determines the reference point.

There are important reasons to understand how these considerations affect estimation of labor supply elasticities. Evaluation of much government policy regarding tax and transfer programs depends on having reliable estimates of the sensitivity of labor supply to wage rates and income levels. To the extent that individuals' levels of labor supply are the result of optimization with reference-dependent preferences, the usual estimates of wage and income elasticities are likely to be misleading.

In this study, I develop an empirical model of daily labor supply that incorporates reference-dependent preferences but does not require that the reference level of income be observed or defined in advance. I apply this model to data on the daily labor supply of New York City taxi drivers by allowing taxi drivers to have a reference level of daily income. The estimates suggest that, while there may be a reference level of income on a given day such that there is a discrete increase in the probability of stopping when that income level is reached, the reference level varies substantially day to day for a particular driver. Additionally, most shifts end before the reference income level is reached. Essentially, the data show more smoothness in the relationship between income and the continuation and stopping probabilities than seems consistent with an important role for reference-dependent preferences.

I. Reference-Dependent Preferences

The theory of reference-dependent preferences has its roots in the concept of loss aversion. Amos Tversky and Daniel Kahneman (1991) present an analysis of choice in a riskless framework where the concavity of utility as a function of income changes at some reference or status quo value of income (Y_R). Marginal utility is higher and increasing in income, at incomes below Y_R. Marginal utility is lower and decreasing in income, at incomes above Y_R. This preference structure implies that, for equivalent absolute dollar gains and losses around Y_R, individuals give up more utility from the dollar loss than they receive from the dollar gain.

This core idea has been credited as the explanation for the "endowment effect" (Richard Thaler 1980). The endowment effect, which has much experimental support, states that individuals

appear to value items they own more than identical items owned by others. For example, an easily replicable experiment is to distribute commemorative mugs to a random subgroup of some experimental subject pool, to elicit from those who received the mugs a willingness to accept (WTA) a given amount in exchange for the mug, and to elicit from those who did not receive the mugs a willingness to pay (WTP) a given amount to purchase a mug. It is a routine finding that the average WTA is substantially greater than the average WTP. Indeed, this relationship between WTA and WTP is considered a test of reference-dependent preferences (Ian Bateman et al. 1997).[1]

While there has been continued refinement of the theoretical ideas underlying reference-dependent preferences (e.g., Alistair Munro and Robert Sugden 2003; Botond Kőszegi and Matthew Rabin 2006), there is relatively little direct field evidence on reference-dependent preferences. One example is David Genesove and Christopher Mayer (2001), who examine seller behavior in the housing market using the purchase price as the reference point. They find clear evidence that sellers are more risk averse when faced with losses. This is consistent with a utility function that changes concavity at the reference point.

There is an emerging literature on labor supply in settings where workers are free to set their labor supply (e.g., Colin Camerer 1997; Gerald S. Oettinger 1999; Ernst Fehr and Lorenz Goette 2007; and Farber 2005).[2] None of this literature allows directly for the possibility of reference-dependent preferences and, by and large, this literature finds substantial positive labor supply elasticities consistent with the standard neoclassical labor supply model.

A notable exception directly relevant to my study is the finding of a negative labor supply elasticity among New York City taxi drivers by Camerer et al. (1997). They interpret their finding as evidence of target earnings behavior by taxi drivers. A target earner is defined as an individual who works until earnings reach the target level, and then quits. Target earnings behavior is an extreme version of reference-dependent preferences, where, at income levels less than the target/reference level, the marginal utility of income is extremely high and, at income levels greater than the target/reference level, the marginal utility of income is extremely low. Essentially, there is such a sharp kink in the utility function at the reference/target level that workers always quit when income reaches the reference/target level. This is clearly nonoptimal from a neoclassical perspective, since it implies quitting early on days when it is easy to make money and working longer on days when it is harder to make money. Utility would be higher by allocating time in precisely the opposite manner. In an earlier paper (Farber 2005), I critiqued the conceptual and econometric approach used by Camerer et al., and I analyzed new data (also used in this study) on the labor supply of New York City taxi drivers using a completely different econometric framework. In that study, I found no evidence of target earnings behavior.

One can also interpret work that considers the effect of relative income or consumption comparisons across individuals on behavior as implicitly using reference-dependent preferences. The idea is that the interpersonal comparisons set up a reference point that introduces an asymmetry between gains and losses relative to this reference point. For example, David Neumark and Andrew Postlewaite (1998) use relative income concerns to understand the increase in labor supply of married women. As more married women enter the labor force and raise their families' incomes, the reference level of income that families use to judge their well-being increases. The result is that more women enter the labor force so that they do not suffer the substantial loss in utility that comes from family income below the reference level.

---

[1] See Kahneman, Jack L. Knetsch, and Thaler (1991) for a nice discussion of the endowment effect and its relationship to reference-dependent preferences.
[2] I survey this literature briefly in Farber (2005).

II. Modeling Labor Supply with Reference-Dependent Preferences

I implement reference-dependent preferences by allowing the marginal utility of income to be higher at income levels below the reference level than at income levels above the reference level. In this case, the utility function is continuous everywhere, but marginal utility is discontinuous at the reference level. This is a natural formulation, given the roots of the reference point concept in the literature on loss aversion.

I show in my earlier work (Farber 2005) that it is not appropriate to characterize the labor supply of a taxi driver as a response to a parametric wage, regardless of the specific utility function. An alternative approach I use is to consider the end of each passenger trip as a decision point regarding whether to continue driving or to stop. In the remainder of this section, I derive the implications of a simple model of reference-dependent preferences for a driver's labor supply decision, assuming that future earnings are stochastic.

One parametrization of a driver's utility (as a function of accumulated income and hours) after trip t during a shift that incorporates a reference point is

$$(1) \qquad U(Y_t, h_t) = (1 + \alpha I[Y_t < \mathrm{T}])(\mathrm{Y} - \mathrm{T}) - \frac{\psi}{1 + \nu} h_t^{1+\nu},$$

where $I[\cdot]$ is the indicator function and

- h_t = hours worked on shift at end of trip t,
- Y_t = income earned on shift at end of trip t,
- T = reference income,
- α = parameter determining sharpness of "kink" in utility function ($\alpha > 0$),
- ψ = parameter determining disutility of work, and
- ν = elasticity parameter (wage elasticity of labor supply = $1/\nu$, defining $Y = Wh$).

This utility function is continuous, but kinked at income level T, with the marginal utility of income discontinuously higher below the kink ($\alpha > 0$) than above the kink.

Next, I express the utility on trip $t + 1$ as a function of accumulated income and hours as of the end of trip t and the fare and time associated with trip $t + 1$. Denote the next fare by f_{t+1} and the time to search for and complete the next fare by τ_{t+1}. On this basis, the utility after trip $t + 1$ can be written as

$$(2) \qquad U(Y_{t+1}, h_{t+1}) = (Y_t + f_{t+1} - T)(1 + \alpha I[Y_t + f_{t+1} < T]) - \frac{\psi}{1 + \nu}(h_t + \tau_{t+1})^{1+\nu}.$$

The marginal utility derived from the next trip t depends on earnings associated with the next fare and the time necessary to complete this trip. For analytic convenience in working with this marginal utility, I approximate the second (hours) term of the utility function using a first-order Taylor series expansion around accumulated hours at the choice point (h_t).[3] I further simplify the expression for marginal utility by assuming that marginal utility on the next trip is at the higher "pre-reference" rate when accumulated income prior to that trip is anywhere below the reference

---

[3] Given the relatively short length of trips, the first-order approximation is likely to be quite good.

income level (even if the next fare causes accumulated income to exceed the reference income level). The resulting expression for marginal utility is

$$(3) \qquad\qquad MU_{t+1} = f_{t+1}(1 + \alpha I[Y_t < \mathrm{T}]) - \tau_{t+1}\psi h_t^\nu.$$

Dividing both sides of equation (3) by the time required for the $t + 1$ trip (τ_{t+1}) yields

$$(4) \qquad\qquad MV_{t+1} = \omega_{t+1}(1 + \alpha I[Y_t < \mathrm{T}]) - \psi h_t^\nu,$$

where MV_{t+1} denotes the marginal utility per unit time on the $t + 1$ trip, which I call "marginal value." The quantity $\omega_{t+1} = f_{t+1}/\tau_{t+1}$ is the wage per unit time on the $t + 1$ trip.

The marginal value of continuing to drive is a linear function of the uncertain wage on the $t + 1$ trip. Let $E(\omega_t) = \mu_t^\omega$ so that expected marginal value is

$$(5) \qquad\qquad E[MV_{t+1}] = \mu_{t+1}^\omega(1 + \alpha I[Y_t < \mathrm{T}]) - \psi h_t^\nu.$$

Conditional on the expected wage, the expected marginal value of continuing to drive declines monotonically as the shift progresses and income and hours worked accumulate. There is a smooth increase in the marginal disutility of hours, as hours accumulate at all hours and income levels. In contrast, the marginal utility of income is constant, aside from the discontinuous drop when the reference income level is reached.[4] If the expected wage is nonincreasing during a shift, then expected marginal value is monotonically declining in t. In this case, the driver would continue to drive until the expected marginal value of the next trip is negative. On this basis, the function relating the continuation probability to accumulated income will exhibit an incremental decline at the reference income level. There should be no other relationship between the continuation probability and accumulated income conditional on a set of observables indicating hours worked and factors affecting expected fares and times (e.g., local, time of day, day of weak, weather).[5]

If, in fact, the expected wage were nonincreasing, I could implement a structural stopping model based on the expected marginal value in equation (5). However, it is unreasonable to assume that the expected wage is nonincreasing. For example, there may be lulls in the middle of a shift where waiting times are rather high. Drivers may not quit for the day at these points because they expect fares to be more plentiful later in the shift. Given that the decision to stop driving for the day is irreversible, consideration of the option value of continuing to drive becomes important. A complete solution to the optimal stopping problem in this context requires the solution of a stochastic dynamic optimization problem that incorporates this option value.[6] Rather than provide this solution (which is not usefully amenable to empirical implementation with the available data), I implement a reduced-form approach to the problem.

---

[4] In the language of dynamic optimization models, the model states that the driver has two state variables, Y_t and h_t. The decision to stop is a smoothly increasing function of h_t but a discontinuous function of Y_t, with a discrete jump in the likelihood of stopping when Y_t reaches T.

[5] This is a result of the constant marginal utility of income assumption built into the specification of the utility function. This assumption is standard in the literature on intertemporal substitution in labor supply, and I discuss its plausibility in more detail in Farber (2005).

[6] This problem arises in other domains, for example, if there is a substantial literature on the retirement decision in the presence of defined-benefit pension plans with nonmonotone marginal values of continuing to work. James H. Stock and David A. Wise (1990) present an analysis of retirement in this context that offers an alternative to a full solution to the problem. See John Rust (1989) and James Berkovec and Steven Stern (1991) for other approaches to estimating dynamic retirement models.

III. An Empirical Model of the Labor Supply Decision

Consider the end of each passenger trip as a decision point regarding whether to continue or to stop driving. A driver will quit for the day when the expected value of continuing to drive first becomes negative. A reasonable approximate solution to the full dynamic stopping problem can be implemented empirically as a discrete choice problem. After any trip t during a shift, a driver can calculate the forward-looking expected value of continuing to drive (the continuation value). This will be a function of many factors, including hours worked so far on the shift and variables that affect expectations about future earnings possibilities. It is also affected by accumulated shift income in an unusual way: when accumulated income is less than the reference income level, there is an opportunity for the driver to derive utility from additional fare income at the higher "pre-reference" rate.

In order the implement the model empirically, I first define a latent variable C_{ijt} for driver i on shift j after trip t that represents the forward-looking expected value of continuing to drive (the continuation value). An empirical representation of this latent variable is

$$(6) \qquad C_{ijt} = X_{ijt}\boldsymbol{\beta} + \delta I\left[Y_{ijt} < T_{ij}\right] + \varepsilon_{ijt},$$

where

- T_{ij} represent the reference income level for driver i on shift j,
- Y_{ijt} represent the income level for driver i on shift j at trip t,
- X_{ijt} is a vector of variables that determine the difference between current utility and the continuation value,
- $\boldsymbol{\beta}$ is a parameter vector to be estimated,
- ε_{ijt} is a random component with a standard normal distribution.
- $I\left[Y_{ijt} < T_{ijt}\right]$ is an indicator function that equals one if accumulated income is less than the reference income level and equals zero otherwise, and
- δ is a positive parameter that represents the increment to the continuation value when the reference income level is greater than income.

The second term in equation (6) embodies the incremental utility value of income below the reference income level that could be earned by continuing to drive.

Driver i on shift j will continue driving after trip t if $C_{ijt} \geq 0$ and stop if $C_{ijt} < 0$. The probability of continuing conditional on T_{ij} is

$$(7) \qquad P^c_{ijt}|T_{ij} = \Phi[X_{ijt}\boldsymbol{\beta} + \delta I\left[Y_{ijt} < T_{ij}\right]],$$

where $\Phi[\cdot]$ is the standard normal cumulative distribution function.

If the reference income (T_{ij}) level were known, this specification would imply the usual probit model, and estimates could be derived by maximum likelihood in the usual way. But the reference income level is not known and needs to be estimated. Suppose the reference income level is

$$(8) \qquad T_{ij} = \theta_i + \mu_{ij},$$

where θ_i is an individual mean reference income level and μ_{ij} is a random component distributed normally with mean 0 and variance σ^2_μ. The μ_{ij} represent daily deviations from θ_i in the reference income level.

The next step is to derive a likelihood function based on the probability that a shift ends (the "shift probability") after trip t in the case where the reference income level is unknown but defined by equation (8). I present a brief description of the likelihood function here.[7] Think of an observation as a shift. If the driver stops after trip t, he did not stop after the first $t-1$ trips. Based on equation (7), it is straightforward to write the joint probability of this pattern conditional on a value for the reference income level (T_{ij}). Given the assumed distribution for the reference income level (equation (8)), I "integrate out" T_{ij} to derive the unconditional probability of driver i on shift j stopping after trip t. The likelihood function for my sample of 538 shifts for 15 drivers is constructed from these unconditional shift probabilities. In Section V, I compute maximum likelihood estimates of the parameters of the model ($\boldsymbol{\beta}$, θ_i, σ_μ^2, and δ).

A. *What Constitutes a Test of Reference Dependence in Labor Supply?*

A central goal of the analysis is to determine whether daily reference income levels are an important factor in the labor supply decisions of taxi drivers. Estimates of the model derived in this section can be used to achieve this goal. The test I use is subjective but relies on a substantive understanding of the potential role of reference-dependent preferences in labor supply.

Reference dependence is important to the extent that it implies predictable behavior changes when income reaches the reference income level. In the context of the model, there is a discrete change in the continuation probability at the reference income level. Empirically, the model implies that drivers stop with higher probability, though not with certainty, when the reference income level is reached. Drivers may continue to work beyond this point on some days if the option value of future fares is sufficiently high.

A test would be straightforward if the reference income level were observed. In this case, one could examine the data for a discrete increase in the stopping probability at the known reference levels. Given that reference income levels are not observed, the test must rely on estimates of the parameters of the econometric model.

The key parameter controlling the change in the continuation probability at the reference point is δ (the increment to the continuation value when income has not reached the reference level). A substantial positive value for δ would imply a significant increase in the probability of stopping when the reference income level is reached.

A large δ is not sufficient for reference-dependent preferences to be useful in predicting labor supply. It must also be the case that a given individual's reference income levels are roughly constant or vary predictably over reasonable periods of time. If this were not the case, then variation across days in labor supply could be rationalized as an optimal response to random changes in the reference level of income. The standard deviation (σ_μ) of the daily random component determining reference income is a measure of how much the reference income level varies from day to day. If the standard deviation is small relative to the mean (θ_i), then I can conclude that reference incomes are fairly stable. In this case, knowing θ_i could provide important information regarding when driver i will stop. Specifically, a driver's mean reference income level (θ_i) will be strongly predictive of his daily income.

Intuitively, if reference dependence is important, it is likely that most shifts will end with drivers reaching their reference income level (though not necessarily stopping at that point). If this is not the case, then reference dependence is not likely to have a substantial effect on behavior. Thus, another test of the model is to examine the predicted probability that the reference income

---

[7] The notation is cumbersome due to the discrete shift in the continuation probability when Y_{ijt} reaches T_{ij}. I present the derivation of the likelihood function in the online Appendix available at http://www.aeaweb.org/articles.php?doi=10.1257/aer.98.3.1069.

level has been reached at shift's end. Given the total income observed on each shift, this is a straightforward calculation based on the estimated parameters of the model.

To summarize, evidence consistent with reference-dependent preferences playing an important role in determining daily labor supply of the taxi drivers in my sample would be:

- A substantial decline in the continuation value when the reference income level is reached (as indexed inversely by the estimated value of δ);

- An estimated value of σ_μ that is small relative to the values estimated for the mean reference income levels (the $\boldsymbol{\theta}$ vector or the mean of the distribution of θ_i);

- A driver's mean reference income level (θ_i) strongly predictive of daily income; and

- Attainment of the reference income level on most shifts in the sample.

I examine each of these pieces of evidence in Section V.

IV. Data

The data necessary to carry out my analysis are available on "trip sheets" that drivers fill out during each shift. Each trip sheet lists the driver's name, hack number, and date, along with details of each trip. The information for each trip includes the start time, start location, end time, end location, and fare. In order to obtain a sample of trip sheets, in the summer of 2000 my research assistants created a list of taxi leasing companies from the current edition of the New York City Yellow Pages. After contacting more than 70 leasing companies, one was found that was willing to provide trip sheets. We were sent 244 trip sheets for 13 drivers covering various dates over the period from June 1999 through May 2000. We contacted the leasing company again in the summer of 2001, and we were sent an additional 349 trip sheets for 10 drivers covering various dates over the period from June 2000 through May 2001. Two of the drivers appear in both groups, so I have a total of 593 trip sheets for 21 drivers over the period from June 1999 through May 2001. A few of these trip sheets refer to common dates, so I have data on 584 shifts. Because I am trying to estimate driver-specific mean reference income levels, I dropped six drivers with ten or fewer shifts from my analysis. These six drivers accounted for a total of 46 shifts, so the analysis sample has a total of 538 shifts and 12,187 trips for 15 drivers.

The drivers in my sample leased their cabs weekly for a fee of $575. Each driver pays for his own fuel and keeps all of his fare income and tips. An unfortunate consequence of receiving the trip sheets in an unsystematic fashion is that I have no information on the number of shifts worked. If a trip sheet is not available for a specific driver on a given day, I cannot determine if that driver did not work on that day or if the trip sheet was simply not provided. This prevents me from examining in any conclusive way inter-day relationships in labor supply.

I performed several regularity checks to insure that the trip sheets were reasonably complete and internally consistent. Where they were not, I cleaned the data using a set of reasonable rules. These rules are outlined in detail in Farber (2005).

Almost all trips (92 percent) started and ended in Manhattan. My earlier work using these data (Farber 2005) suggests that there are important differences in stopping behavior between trips ending outside Manhattan and trips ending within Manhattan, but no significant differences within these two broad locations. I proceed using an indicator to distinguish Manhattan from other locations.

TABLE 1—SHIFT LEVEL SUMMARY STATISTICS, BY DRIVER

Driver	Number of shifts	Average trips	Working hours	Driving hours	Waiting hours	Break hours	Total income	Average wage
Driver 1	39	23.56	6.85	4.32	2.53	0.90	157.58	23.16
Driver 2	14	12.29	3.89	2.78	1.11	2.41	97.10	25.11
Driver 3	40	22.10	6.28	4.52	1.76	0.39	147.51	23.89
Driver 4	23	16.52	6.46	3.98	2.48	2.11	144.96	23.65
Driver 5	24	22.29	6.47	4.42	2.05	0.74	160.71	25.59
Driver 6	37	25.32	7.78	5.13	2.64	0.86	172.44	22.54
Driver 7	19	25.58	7.17	5.47	1.70	0.54	162.02	23.23
Driver 8	45	20.27	6.35	3.90	2.45	1.65	133.19	21.46
Driver 9	13	19.46	6.15	4.03	2.13	0.55	157.95	25.78
Driver 10	17	21.29	7.06	4.49	2.57	0.64	165.84	23.59
Driver 11	70	25.06	6.84	4.56	2.28	0.93	172.01	25.62
Driver 12	72	28.10	8.53	5.84	2.69	0.60	203.05	24.01
Driver 13	33	17.06	6.91	4.63	2.29	0.97	163.51	23.91
Driver 14	46	24.46	7.10	4.80	2.30	0.67	156.23	22.00
Driver 15	46	19.17	5.32	3.66	1.66	0.24	128.97	24.72
All	538	22.65	6.84	4.58	2.26	0.87	160.03	23.80

I additionally collected data from the National Atmospheric and Oceanic Administration (NOAA) on temperature and precipitation in New York City. I collected daily average, minimum and maximum temperatures, and total daily rainfall and snowfall in Central Park. I also collected hourly rainfall data at LaGuardia Airport.

A. *Summary Statistics*

Table 1 contains average statistics by shift for each driver. Hours worked per day is defined as the sum of driving time (the sum over trips of the time between the trip start time and the trip end time) and waiting time (the sum over trips of the time between the end of the last trip and the start of the current trip). Waiting time is substantial, accounting for 33 percent of working time on average. Break time averages about 52 minutes per shift.

There is substantial variation across drivers in average hours worked per day, with means ranging from 3.89 to 8.62. Still, the majority of the variation in daily work hours is within-driver variation across days. The standard deviation of daily work hours is 2.50. The R-squared from a regression of daily hours on a set of driver fixed effects is 0.152, with a residual root mean squared error of 2.33. Panel A of Figure 1 contains a histogram of hours worked for the 538 shifts. The distribution is single-peaked, with the mode at eight hours.

There is also substantial variation across drivers in total fare income per day, with means ranging from $97.10 to $228.26[8]. Not surprisingly, daily income covaries strongly with daily hours with a simple correlation of 0.91. As with hours, the majority of the variation in daily income is within-driver variation across days. The standard deviation of daily income is $59.82. The R-squared from a regression of daily income on a set of driver fixed effects is 0.153 with a residual root mean squared error of $55.77. Panel B of Figure 1 contains a kernel density estimate of daily income.[9]

---

[8] Income per day is the sum of fares. Tip income is not measured or accounted for.
[9] All kernel density estimates in this study use the Epanechnikov kernel. The bandwidths are listed in the figures.
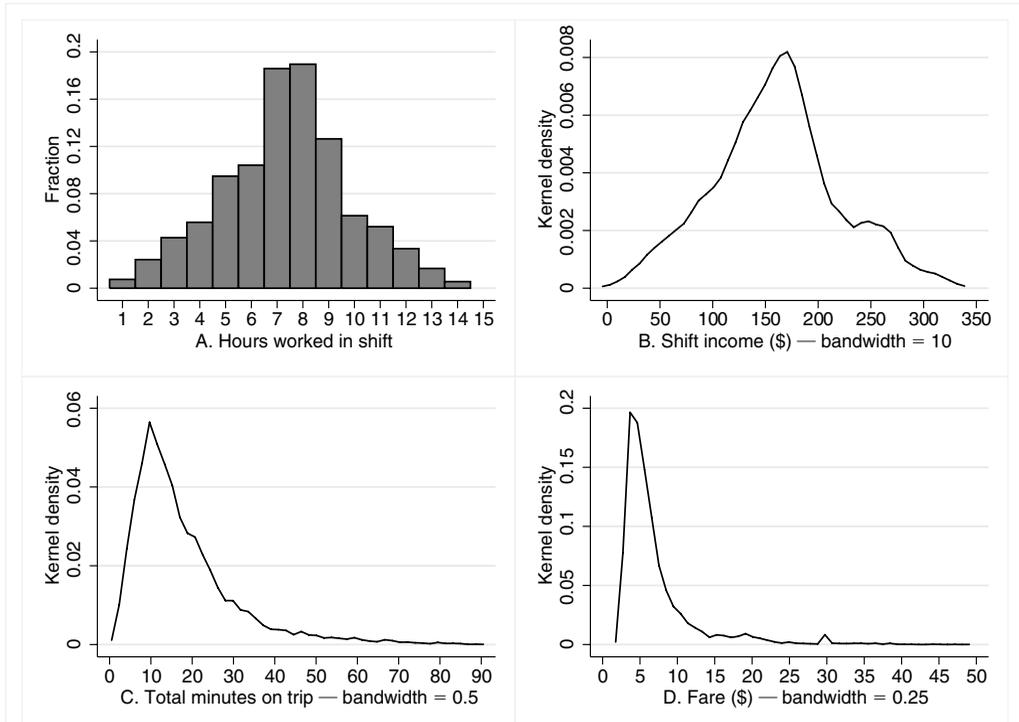
FIGURE 1. DISTRIBUTIONS OF HOURS AND INCOME, BY SHIFT AND KERNEL DENSITY ESTIMATES, OF TOTAL TRIP TIMES AND FARES

The last column of Table 1 contains the daily average for each driver of their hourly wage rate (total income divided by working hours). These show less inter-driver variation, ranging from a low of $21.46 to a high of $25.78. The standard deviation of the daily wage rate is $4.52. Most of this is within-driver variation, as the R-squared from a regression of the daily wage on a set of driver fixed effects is 0.082 with a residual root mean squared error of $4.39.

Panel C of Figure 1 contains a kernel density estimate of the distribution of total trip times (the sum of waiting and travel times) for the 12,187 trips in my sample.[10] Median time per trip is 15 minutes and the mean is 18.1 minutes. That trips are this short reflects the fact that 92.5 percent of the trips in my sample begin and end in Manhattan. The fact that the great majority of trips are short implies that opportunities to stop driving occur frequently.

Panel D of Figure 1 contains a kernel density estimate of the distribution of fares.[11] The median fare is $5.30, and the mean fare is $7.06. Once again, the small size of the fares (which exclude tips) is due to the fact that most trips are intra-Manhattan. The average intra-Manhattan fare is $5.93, while the average fare that starts or ends outside Manhattan is $20.76. The small blip at $30 represents the flat rate between Kennedy Airport and Manhattan in force during my sample period. The fact that the great majority of fares are small implies that the first opportunity to stop driving after the reference income is exceeded will generally arise at an income level not greatly in excess of the reference income level.

---

[10] I have truncated this distribution at 90 minutes for total trip time. There are 38 trips with times greater than 90 minutes.

[11] I have truncated this distribution at $50 for fares. There are eight trips with fares greater than $50.

TABLE 2—MAXIMUM LIKELIHOOD ESTIMATES OF REFERENCE-DEPENDENT LABOR SUPPLY MODEL

Parameter	(1)	(2)	(3)	(4)
$\hat{\beta}$ (contprob) (constant)	−0.691	—	—	—
	(0.243)			
$\hat{\theta}$ (mean ref inc)	159.02	206.71	250.86	—
	(4.99)	(7.98)	(16.47)	
$\hat{\delta}$ (cont increment)	3.40	5.35	4.85	5.38
	(0.279)	(0.573)	(0.711)	(0.545)
$\hat{\sigma}^2$ (ref inc var)	3,199.4	10,440.0	15,944.3	8,236.2
	(294.0)	(1,660.7)	(3,652.1)	(1,222.2)
Driver $\hat{\theta}_i$ (15)	No	No	No	Yes
Vars in cont prob				
Driver FE's (14)	No	No	Yes	No
Accum hours (7)	No	Yes	Yes	Yes
Weather (4)	No	Yes	Yes	Yes
Day shift and end (2)	No	Yes	Yes	Yes
Location (1)	No	Yes	Yes	Yes
Day-of-Week (6)	No	Yes	Yes	Yes
Hour-of-Day (18)	No	Yes	Yes	Yes
Log(L)	−1,867.8	−1,631.6	−1,572.8	−1,606.0
Number of parameters	4	43	57	57

Note: The sample includes 12,187 trips in 538 shifts for 15 drivers. Models 2–4 also include a constant in the *X* vector for the continuation probability. The hours from 5 AM to 10 AM have a common fixed effect. Standard errors are reported in parentheses.

V. Estimation of the Labor Supply Model

In this section I present maximum likelihood estimates of the parameters ($\boldsymbol{\beta}$, θ_i, σ_μ^2, and δ) of the reference-dependent labor supply model based on the likelihood function described in Section III and on the data described in the previous section.

The first column of Table 2 contains estimates of a restricted model that constrains all drivers to have the same mean reference income level and the *X* vector to contain only a constant. The evidence is mixed with regard to the role of reference income levels in determining labor supply. The estimated mean reference income level is reasonable at $159.02. The estimate of δ, which indexes (inversely) the change in the probability of continuing to drive once the reference income level for the day is reached, is substantial at 3.40 and statistically significantly different from zero. This would seem to be strong evidence of the importance of reference-dependent preferences in this context.

There is substantial inter-shift variation, however, around the mean reference income level. The estimated variance of 3,199.4 implies a standard deviation of $56.60. To the extent that this represents daily variation in the reference income level for a particular driver, the predictive power of the reference income level for daily labor supply would be quite limited.

The second column of Table 2 contains estimates of the model that include a set of variables in the continuation function (equation (6)) that are meant to capture earnings opportunities and other factors that would affect a driver's continuation probability. These are:

- Indicators for eight categories of hours worked at trip end;
- Six indicators for the day of week;
- Indicators for 18 clock hours at trip end;
- A day-shift indicator;

- An interaction of day-shift with clock hours 3–4 PM meant to capture the likelihood that a day-shift driver must turn the car over to a night-shift driver at that time of day;
- Four variables measuring weather, including daily snowfall, hourly rainfall, high heat (maximum temperature $>=$ 80 degrees), and cold (minimum temperature $<$ 30 degrees); and
- An indicator for non-Manhattan location.

These 39 variables are clearly important in determining labor supply as the log-likelihood improves from $-1{,}867.8$ to $-1{,}631.6$ when they are included in the model. The estimated mean reference income level increases to \$206.71. The estimate of δ increases to 5.35, implying once again that reference income factors may be important, but the inter-shift variance of reference income around its mean increases dramatically to 10,440.0 implying a standard deviation in the reference income level of \$102.17.

Column 3 of Table 2 contains estimates of the model where the continuation function additionally includes driver fixed effects. These capture inter-driver differences in the marginal disutility of labor. The driver fixed effects dramatically improve the fit of the model, increasing the log-likelihood from $-1{,}631.6$ to $-1{,}572.8$. The estimates of the reference-dependence parameters continue to be characterized by (a) a substantial increment (δ) to the continuation function before the reference income level is reached; and (b) substantial variation in the reference income level.

The first three columns of Table 2 show an important constraint in the models, that is, all drivers are assumed to have the same mean reference income level. Indeed, the large estimate of the variance in the reference income level could be due to differences across drivers in their mean reference income level. The obvious solution is to estimate separate mean reference incomes for each driver. However, estimation of two sets of driver fixed effects in the model is asking too much of the data.[12] As a result, I estimate a model with unique mean reference income levels for each driver (θ_i), but without driver fixed effects in the continuation function. These estimates are contained in column 4 of Table 2.

Allowing for different mean reference income levels improves the fit of the model substantially, with the log-likelihood improving from $-1{,}631.6$ (in column 2 of the table) to $-1{,}606.0$. The hypothesis that all drivers that have the same mean reference income level can be rejected at any reasonable significance level (p-value $= 3.8$E-6) using a likelihood-ratio test. Interestingly, the model with fixed effects in the continuation function (column 3) fits the data better (a higher log-likelihood value) than does the model with driver-specific mean reference incomes.[13]

With regard to estimates of parameters related to the role of reference income levels on labor supply, the evidence remains mixed. The general pattern across the specifications in Table 2 is similar, and I focus the discussion on the estimates in column 4. The estimate of δ, which indexes (inversely) the change in the probability of continuing to drive once the reference income level for the day is reached, is substantial at 5.38 and significantly different from zero (p-value $< 1.$E-10). Given the standard normal distribution of the latent variable determining the probability of continuing to drive, as shown in equation (6), this implies that the continuation probability is close to zero once the reference income level is reached, regardless of when in the shift is reached. This would seem to be strong evidence of the importance of reference-dependent

---

[12] I did attempt to estimate such a model, and I was not able to find an optimum, despite considerable effort. It is likely that more observations (both in the form of more drivers and more shifts per driver) are necessary to estimate such a model.

[13] While I was not able find a unique optimum for the full model that nests the two models in columns 3 and 4 of Table 2, I was able to find log-likelihood values in the range of $-1{,}554.4$. This is sufficiently high to be able to reject that either of the sets of fixed effects is zero. Thus, it appears that mean reference income levels do vary by driver even after driver differences in the continuation function are accounted for.

TABLE 3—MAXIMUM LIKELIHOOD ESTIMATES OF
REFERENCE-DEPENDENT LABOR SUPPLY MODEL
(*Column 4 of Table 2*)

Variable	$\hat{\beta}$	Driver	$\hat{\theta}$
Hour ≤ 2	2.242	1	245.66
	(0.601)		(28.27)
Hour 3–5	1.420	2	93.06
	(0.248)		(37.87)
Hour 6	0.446	3	166.95
	(0.153)		(20.97)
Hour 7	0.210	4	181.83
	(0.131)		(31.72)
Hour 9	0.094	5	185.18
	(0.187)		(21.63)
Hour 10	−0.601	6	284.67
	(0.210)		(40.53)
Hour 11	−0.763	7	211.44
	(0.258)		(35.28)
Hour ≥ 12	−1.088	8	190.40
	(0.327)		(18.76)
Min temp < 30	−0.033	9	182.52
	(0.130)		(39.13)
Max temp ≥ 80	0.003	10	186.61
	(0.153)		(32.00)
Hourly rain	−0.225	11	217.17
	(0.178)		(16.13)
Daily snow	0.035	12	231.93
	(0.171)		(12.80)
Non-Manhattan	−0.568	13	181.39
	(0.166)		(27.14)
		14	232.82
			(23.79)
		15	150.82
			(21.51)

Notes: The sample includes 12,187 trips in 538 shifts for 15 drivers. The model also includes in the continuation function a constant, 6 indicators for day of the week, a day-shift and end-of-day shift indicator, and 18 indicators for hour of the day. The hours from 5 AM to 10 AM have a common fixed effect. Standard errors are reported in parentheses. See column 4 of Table 2 for more details.

preferences in this context. While drivers may or may not stop before their reference income level is reached, they are almost sure to stop once they reach that level.

With the allowance for driver-specific mean reference levels, the variance of the reference income level reflects within-driver day-to-day variation in reference income. What is perhaps surprising is that the estimate of the within-driver variance remains very large at 8326.2, implying a within-driver daily standard deviation in reference income of $91.25. Based on the assumption of normality, the inter-quartile range for a particular driver with a mean reference income level of $250 is from $188 to $311, and the 95 percent confidence interval is from $99 to $400. The large variances estimated in the first three columns of Table 2 are not simply the result of different drivers having different mean reference income levels.

Table 3 contains more detailed estimates of the full model in column 4 of Table 2. The estimates of β, in the first column, show a strongly decreasing continuation probability with hours worked on the shift. The standard normal latent variable determining the probability of continuing (equation (6)) decreases by 3.3 from the beginning of the shift to the twelfth hour. The weather does not appear to be significantly related to the labor supply decision. Ending a fare outside of Manhattan results in a significantly lower probability of continuing, probably due to the fact that many drivers live outside of Manhattan.

The estimates of the individual mean reference income levels (the θ vector) vary significantly across drivers, from a low of $93.06 for driver 2 to a high of $284.67 for driver 6. The mean reference income is $196.16 with a standard deviation across the 15 drivers of $44.70.

VI. Does Reference Dependence Have Predictive Value?

In Section IIIA, I listed four types of evidence consistent with reference dependent preferences playing an important role in determining daily labor supply.

1. It is clearly the case that the continuation probability decreases dramatically when the reference income level is reached. The estimated value of δ of 5.38 (column 4 of Table 2) implies a decrease of over five standard deviations in the standard normal distribution of the continuation value when the reference income level is reached. It appears that drivers are almost certain to stop after they reach their reference income level for the day.

2. The large estimated within-driver inter-day variance in the reference income level implies that there is little consistency in a driver's reference income level day to day.

3. There is a strong positive correlation between both average income and average hours across shifts for a given driver and that driver's estimated mean reference income level (θ_i). These correlations are 0.76 and 0.86, respectively. In order to investigate whether θ_i is a strong predictor of daily income, I estimate the following simple regression at the shift level of income on the estimated mean reference income level for the 538 shifts in my sample:

$$(9) \qquad\qquad Y_{ij} = \alpha_0 + \alpha_1 \hat{\theta}_i + \gamma_{ij},$$

where Y_{ij} is income for driver i on shift j. While the estimate of α_1 is significantly positive at 0.408 (s.e. = 0.063), it is substantially less than one and the R^2 from this regression is only 0.07. Thus, very little of the overall variation in income across shifts can be accounted for by variation across drivers in the estimated mean reference income level.[14]

4. The probability that the reference income level is attained on a given shift is relatively low. The probability that income on shift j for driver i (Y_{ij}) exceeds the reference level for that shift is the CDF of a normal distribution with mean θ_i and variance equal to 8,236.2. This is:

$$(10) \qquad\qquad P(T_{ij} \le Y_{ij}) = \Phi[(Y_{ij} - \theta_i)/\sigma_\mu].$$

Figure 2 contains a plot of the fraction of shifts where the predicted probability that the reference income level was reached during the shift falls in each of ten intervals from zero to one. It is clear that the reference income level is reached in only a minority of cases. The mean probability is 0.34, implying that the reference income level was reached in only 34 percent of shifts in the sample. The seventy-fifth percentile of the distribution of probabilities that the reference income level is reached is 0.49, and the ninetieth percentile of the same distribution is 0.63. Thus, it appears that drivers generally stop before the reference income level is reached.[15]

Taken together, these findings leave a puzzle. The reference-dependent utility model builds a particular nonlinearity into the relationship of income with the continuation probability, and the large estimated value of δ is evidence for this nonlinearity: drivers generally will stop if they reach their reference income level. The reference income level for a given driver varies unpredictably day to day, however, and most shifts end before the driver reaches his reference income level. Essentially, the data show more smoothness in the relationship between income and the continuation and stopping probabilities than seems consistent with an important role for reference-dependent preferences. The relatively large unexplained within-driver variation in income across shifts also belies an important role for daily reference dependence in this context.

---

[14] In fact, most of the variation in income across shifts is within-driver variation. A regression of daily income on a set of driver fixed effects yields an R^2 of 0.15.

[15] In order to investigate the possibility that drivers end their shift when they are near, but not necessarily at, the reference income level, I repeat this analysis calculating for each shift the probability that, at shift's end, accumulated income exceeds 90 percent of the reference income level. While these probabilities are higher, they are still fairly low. The mean probability that a shift ends at some point after 90 percent of the reference income level has been reached is 0.41, implying that 90 percent of the reference income level was reached in only 41 percent of shifts in the sample.
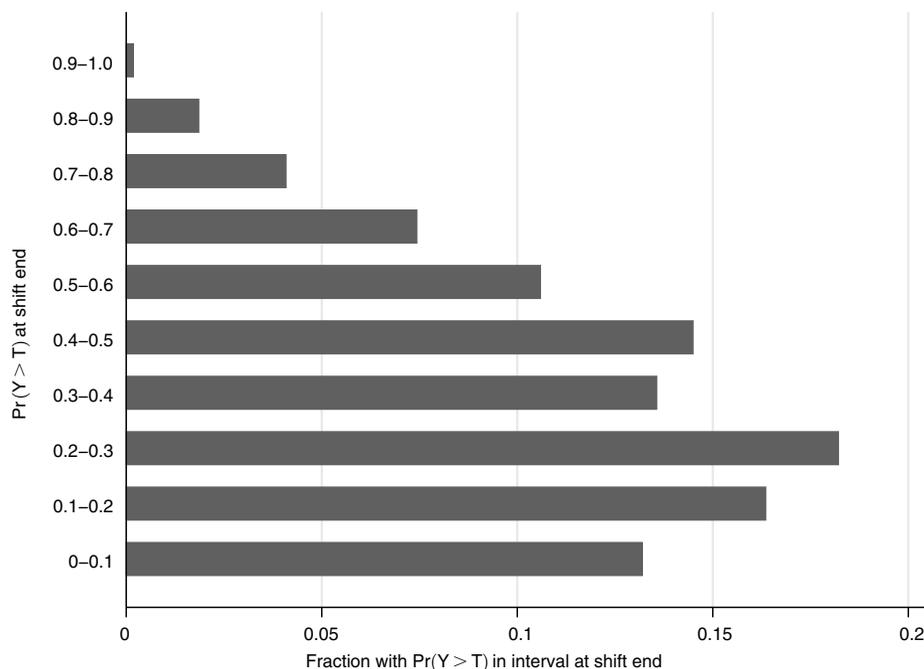
Figure 2. Distribution of Predicted Probability of Income Exceeding the Reference Level at Shift End
(Based on parameters in column 4 of Table 2)

REFERENCES

▶**Bateman, Ian J., Alistair Munro, Bruce Rhodes, Chris Starmer, and Robert Sugden.** 1997. "A Test of the Theory of Reference-Dependent Preferences." *Quarterly Journal of Economics*, 112(2): 479–505.

▶**Berkovec, James, and Steven Stern.** 1991. "Job Exit Behavior of Older Men." *Econometrica*, 59(1): 189–210.

▶**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.

▶**Farber, Henry S.** 2005. "Is Tomorrow Another Day? The Labor Supply of New York City Cabdrivers." *Journal of Political Economy*, 113(1): 46–82.

▶**Fehr, Ernst, and Lorenz Goette.** 2007. "Do Workers Work More if Wages Are High? Evidence from a Randomized Field Experiment." *American Economic Review*, 97(1): 298–317.

▶**Genesove, David, and Christopher Mayer.** 2001. "Loss Aversion and Seller Behavior: Evidence from the Housing Market." *Quarterly Journal of Economics*, 116(4): 1233–60.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler. 1991. "The Endowment Effect, Loss Aversion, and Status Quo Bias: Anomalies." *Journal of Economic Perspectives*, 5(1): 193–206.

▶**Kőszegi, Botond, and Matthew Rabin.** 2006. "A Model of Reference-Dependent Preferences." *Quarterly Journal of Economics*, 121(4): 1133–65.

▶**Munro, Alistair, and Robert Sugden.** 2003. "On the Theory of Reference-Dependent Preferences." *Journal of Economic Behavior and Organization*, 50(4): 407–28.

▶**Neumark, David, and Andrew Postlewaite.** 1998. "Relative Income Concerns and the Rise in Married Women's Employment." *Journal of Public Economics*, 70(1): 157–83.

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

Rust, John. 1989. "A Dynamic Programming Model of Retirement Behavior." In *The Economics of Aging*, ed. David A. Wise, 359–98. Chicago: University of Chicago Press.

▶**Stock, James H., and David A. Wise.** 1990. "Pensions, the Option Value of Work, and Retirement." *Econometrica*, 58(5): 1151–80.

▶**Thaler, Richard.** 1980. "Toward a Positive Theory of Consumer Choice." *Journal of Economic Behavior and Organization,* 1(1): 39–60.

▶**Tversky, Amos, and Daniel Kahneman.** 1991. "Loss Aversion in Riskless Choice: A Reference-Dependent Model." *Quarterly Journal of Economics*, 106(4): 1039–61.