Daily Labor Supply and Adaptive Reference Points

By Neil Thakral and Linh T. Tô*

This paper provides field evidence on how reference points adjust, a degree of freedom in reference-dependence models. Examining this in the context of cabdrivers' daily labor-supply behavior, we ask how the within-day timing of earnings affects decisions. Drivers work less in response to higher accumulated income, with a strong effect for recent earnings that gradually diminishes for earlier earnings. We estimate a structural model in which drivers work towards a reference point that adjusts to deviations from expected earnings with a lag. This dynamic view of reference dependence reconciles conflicting "neoclassical" and "behavioral" interpretations of evidence on daily labor-supply decisions.

In the classical economic model of labor supply, individuals choose hours of work to trade off the utility of additional income against the disutility of additional effort. Based on an analysis of daily decisions about work hours among New York City (NYC) taxi drivers, Camerer et al. (1997) propose the alternative hypothesis that drivers quit working upon reaching a target level of earnings. An ongoing debate since then focuses on the question of whether workers exhibit such reference-dependent behavior with respect to daily earnings.

The broader question of what determines the reference point poses a challenge for evaluating the importance of reference dependence in any given setting. The highly influential work on prospect theory by Kahneman and Tversky (1979) describes the implications of the existence of a reference point but leaves the reference point itself largely unspecified. In an attempt to discipline the theory, Kőszegi and Rabin (2006) endogenize the reference point through assuming that it coincides with recent expectations. Even under this perspective, there remains an implicit degree of freedom in the theory—the speed of adjustment of the reference point—which can substantially affect its empirical predictions.

Understanding whether labor-supply behavior exhibits reference dependence requires both detailed data and an operational model of what the reference point is. This paper uses comprehensive trip-level data on all NYC cab fares in 2013 to identify the timing of reference-point effects. We find that the key income-related determinant of the decision to stop working is recent earnings, not daily earnings,

* Thakral: Department of Economics, Brown University, Box B, Providence, RI 02912, (email: neil_thakral@brown.edu). Tô: Department of Economics, Harvard University, Littauer Center, 1805 Cambridge Street, Cambridge, MA 02138, (email: linhto@fas.harvard.edu). We are grateful to Isaiah Andrews, Pedro Bordalo, Stefano DellaVigna, Henry Farber, Josh Feng, Benjamin Golub, Jerry Green, Kareem Haggag, Maximilian Kasy, Lawrence Katz, Marc Kaufmann, David Laibson, Adam McCloskey, Ted O'Donoghue, Matthew Rabin, Joshua Schwartzstein, Jesse Shapiro, Andrei Shleifer, and Elie Tamer for advice and suggestions.

and the reference-point effects gradually diminish for earnings accumulated earlier in the day. To interpret the results, we offer a conceptual framework which emphasizes the role of reference-point adaptation and structurally estimate the model. By characterizing the dynamics of reference points in the context of daily labor-supply decisions, this paper provides field evidence on a degree of freedom in one of the central models of behavioral economics (DellaVigna, 2009). The framework also helps to organize, explain, and reconcile conflicting interpretations of previous evidence.

We present a model of reference-dependent preferences with an adaptive reference point. The formalization provides a way of capturing the following intuition which conveys our main findings: people overreact to surprises, as they work less in response to higher accumulated earnings, but surprises wear out over time, so that quitting depends to a greater extent on more recent earnings. Our formulation nests the neoclassical model as well as static income targeting. At one extreme lies a reference point that adjusts instantaneously, which produces behavior that coincides with the neoclassical prediction, and at the other extreme lies a fixed reference point. Neither extreme case permits stronger reactions to more recent experiences. In the intermediate case, the reference point exhibits some degree of persistence or stickiness, with decreasing weights on lagged values of the reference point. This slow-adjusting reference point incorporates earlier earnings to a greater extent than more recent earnings, thus accounting for the gradually diminishing effect of earnings on quitting.

To identify the timing of reference-point effects, we use data consisting of over 170 million trips from over 40,000 NYC cabdrivers to estimate how income accumulated at different times influences the decision to stop working. Isolating variation in earnings requires flexibly controlling for factors influencing the value of quitting, including cumulative daily work hours. The data enable us to use a flexible estimation approach to obtain, as a function of how long a driver has been working, the marginal effect of earnings accumulated during different hours of a shift on the decision to quit. An empirical Monte Carlo exercise validates our estimation approach.

We find strong income effects for recent earnings—despite the fact that higher recent earnings predict better opportunities from continuing to work—and behavior that appears neoclassical in response to earnings accumulated earlier in the day. Overall, for a driver who finishes a trip after 8.5 hours of work, a 10 percent increase in accumulated daily earnings corresponds to a 3 percent increase in the probability that he stops working for the day. The effect size changes to 10 percent if the additional earnings come in the most recent hour and gradually declines for earnings accumulated earlier. Earnings from the first four hours have little or no effect on the decision of whether to end a shift at 8.5 hours.

To quantify the speed of adaptation of the reference point, we estimate a structural model of daily labor supply. The model enables us to examine alternative specifications of the reference point in more detail. Crawford and Meng (2011) provide evidence that drivers' expectations of daily earnings as determined by past outcomes serve as the reference point. The adaptive reference point we propose reduces to their beginning-of-day expectation of daily earnings in the special case that the reference point does not adjust within the day. Maximum likelihood estimates of the model confirm that an intermediate degree of adaptation provides a better fit for the patterns in the data: the reference point adjusts to incorporate about 40 percent of a shock to earnings within an hour and about 90 percent within four hours. As an alternative specification of the reference point that also accommodates within-day updating, we consider forward-looking reference points based on a one-period lag of expectations (Kőszegi and Rabin, 2006, 2009) under different definitions of the lag (e.g., previous trip instead of previous day). Adopting a discrete view of how the reference point adjusts would produce a stark contrast between the most recently accumulated earnings, which the reference point does not incorporate, and any earlier earnings, which the reference point does incorporate. The data instead show a gradual decline in the influence of less recent earnings on stopping decisions, and we find that earlier lags remain important for explaining the patterns we observe. In this sense, one can view our framework as extending the Kőszegi and Rabin model by allowing for a gradual adjustment of the reference point.

To put our findings in perspective, we review the previous work on reference dependence in labor supply. The earliest work focuses on estimating daily wage elasticities on the intensive margin for cabdrivers, uncovering a negative relationship between average wages and the number of hours worked each day (Camerer et al., 1997; Chou, 2002). To explain the puzzling finding of a backward-bending labor-supply curve, Camerer et al. (1997) argue that a cabdriver's marginal utility of income must drop sharply around the level of average daily income due to loss aversion, resulting in a probability of quitting for the day that rises substantially when a driver gets near their target.[1] Due to econometric problems with estimating daily wage elasticities, Farber (2005) tests for reference dependence by using a hazard specification to examine directly whether stopping decisions respond to accumulated daily earnings.[2] Despite finding a positive association between accumulated daily earnings and the probability of ending a shift conditional on hours worked, qualitatively consistent with reference dependence, Farber cannot

[1]Chou (2002) lacks data on hourly wages but points out that with such data, "income targeting may be tested more rigorously... utilizing a hazard specification" in which "the probability that a driver quits for the day at any point in time may be parameterized as a function of the cumulated income and the expected marginal wage... Short-horizon targeting predicts that quitting is related to cumulative same-day income..."

[2]Camerer et al. (1997) instrument for a driver's average wage on a given day with summary statistics of the distribution of other drivers' wages on the same day to address potential concerns about division bias (i.e., that average hourly wages obtain from dividing daily income by hours, so measurement error in hours can lead to a spurious negative relationship between wages and hours). However, as Farber (2005) points out, the instrumental-variables approach does not purge the elasticity estimates of day-specific factors that affect both wages and aggregate labor supply. In addition, as Camerer et al. acknowledge, estimating a daily wage elasticity requires that wages vary across days but remain relatively constant within days; Farber (2005) disputes the premise of relatively constant within-day wages. We also document how within-day variation in wages can lead to biased estimates (see Appendix B).

reject the neoclassical null hypothesis with these data.[3]

The availability of large-scale administrative data since then provides an opportunity to settle these unresolved issues. Farber (2015) uses a sample of 13 percent of all NYC cabdrivers between 2009 and 2013 to revisit the earlier studies and test a model of reference dependence with a fixed daily income target. Applying the previous approaches to the comprehensive new dataset leaves a puzzle, as some evidence of income-targeting behavior emerges. Farber (2015) finds that negative wage elasticities appear for one-third of day-shift drivers and one-seventh of night-shift drivers, and accumulated daily income has a small but statistically significant influence on the decision to quit working during day shifts.[4] The modest evidence for the presence of reference dependence raises the possibility of misspecification of the reference point as a potential explanation, especially in light of theoretical work describing the reference point as "recent expectations" (Kőszegi and Rabin, 2006, 2009).

A dynamic view of reference dependence helps to resolve some of the conflicting perspectives in the literature on labor supply. Existing work on daily labor supply using observational and experimental data in a variety of settings yields mixed results.[5] Previous research implicitly tends to offer a binary characterization of behavior, with a negative wage elasticity corresponding to daily income targeting and a positive wage elasticity corresponding to the neoclassical model, or a positive marginal effect of accumulated earnings on stopping corresponding to daily income targeting and a null effect corresponding to the neoclassical model. By emphasizing reference-point adaptation, our model describes the extent to which workers exhibit neoclassical behavior through the speed of adjustment. Our results point towards the relevance of a component of the utility function pertaining to recent expectations in addition to all the forces embedded in the neoclassical model. Neglecting the importance of recency in forming the reference point can lead to misspecification in tests of reference dependence. For example, small effects of accumulated daily income on cabdrivers' quitting decisions (Farber, 2005, 2015) and substantial variation across shifts in their estimated reference income levels (Farber, 2008) would suggest a limited role for reference dependence under

[3]Farber (2008) estimates a structural version of the stopping model, which consists of a latent underlying distribution of daily income targets and accommodates a threshold effect of exceeding the income target, and concludes that the variation in driver-day targets leaves the model with little predictive value despite finding a threshold effect. Crawford and Meng (2011) use the same data to estimate a structural stopping model that allows for reference dependence in both daily income and hours, where drivers' expectations of daily income and hours based on previous shifts determine the driver-day targets following the ideas from Kőszegi and Rabin (2006), and conclude that the data support this model.

[4]Morgul and Ozbay (2014) use the full set of over 30,000 drivers in four separate months of 2013 to revisit the earlier studies as well, finding a negative wage elasticity for the month of January as well as a positive relationship between daily earnings and stopping conditional on hours during all four months.

[5]Papers using observational data that find negative wage elasticities include Ashenfelter, Doran and Schaller (2010), Doran (2014), and Schmidt (2017) on taxi drivers in NYC; Chang and Gross (2014) on pear packers in California; Agarwal et al. (2015) on taxi drivers in Singapore; and Nguyen and Leung (2015) on swordfish fishermen in Hawaii. Those finding positive wage elasticities include Jonason and Wållgren (2013) on taxi drivers in Stockholm and Stafford (2015) on lobster fishermen in Florida. Also see field experiments by Fehr and Goette (2007), Andersen et al. (2018), and Dupas, Robinson and Saavedra (2019).

the assumption of a daily-level income target as the reference point.[6]

Our investigation of labor supply also reveals lessons about models of reference dependence. While some lab experiments find evidence of forward-looking expectations-based reference points following Kőszegi and Rabin (2006), other recent experiments yield mixed results.[7] Existing empirical tests of reference dependence tend to assume a particular view of what constitutes the reference point, including how quickly the reference point adapts to experimental manipulations in the context of lab studies. Motivated by this observation, Heffetz (2018) argues that changing the reference point requires a "sense of internalization" of updated expectations, related to the idea that realization of gains or losses leads to reference-point updating (Imas, 2016).[8] Our paper contributes to this line of work by characterizing the precise timing of reference-point effects. We evaluate a reference point based on a one-period lag of expectations but find support for further history dependence in modeling the reference point. The results complement existing evidence on the importance of past outcomes in shaping the reference point in various domains such as housing demand (Simonsohn and Loewenstein, 2006), risky choice (Post et al., 2008), and job search (DellaVigna et al., 2017), which we return to in the concluding section.

The paper proceeds as follows. The next section provides background information on the institutional context and describes the data. Section II analyzes the impact of accumulated daily earnings on labor supply and discusses possible explanations for the results. Section III presents a model of loss aversion with an adaptive reference point along with structural estimates. Section IV concludes.

I. Data

A. *Background*

Our study uses trip-level data provided by the NYC Taxi and Limousine Commission (TLC) for every fare served by NYC medallion taxicabs in 2013. The "trip sheets" consist of detailed information about each fare, including anonymized identification numbers for the driver and car, start and end times for each trip, pick-up and drop-off locations, tips paid by credit card, and the fare charged. These data are collected and transmitted electronically in accordance with the Taxicab Passenger Enhancements Project (TPEP). Haggag and Paci (2014) and Farber (2015) provide further details about the data, with the former using data from 2009 and the latter using data from 2009–2013.

[6]Similarly, designing a field experiment to test reference dependence by inducing unexpected cash windfalls in the morning (Andersen et al., 2018) makes use of the assumption of a daily-level income target, but a null effect on aggregate labor supply would not reject expectations-based reference dependence if the reference level adjusts during the day.

[7]See Abeler et al. (2011), Ericson and Fuster (2011), Banerji and Gupta (2014), Karle, Kirchsteiger and Peitz (2015), and Sprenger (2015); but also see Heffetz and List (2014), Gneezy et al. (2017), and Cerulli-Harms, Goette and Sprenger (2019).

[8]Some recent lab evidence also explicitly considers the speed of reference-point adjustment (Gill and Prowse, 2012; Song, 2016; Baucells, Weber and Welfens, 2011).

Prior to TPEP, cabdrivers were required to fill out trip sheets by hand to record and store information on paper about each fare. By 2008, all medallion taxicabs in NYC had implemented a series of technology-based service improvements (e.g., credit/debit card payment systems, passenger information monitors, and text messaging between the TLC and drivers) due to a March 2004 mandate by the TLC Board of Commissioners, which also led to automated trip sheet data collection. Relative to the earlier handwritten trip sheets, the electronically transmitted data also include Global Positioning System (GPS) coordinates for pick-up and drop-off locations, available for over 98 percent of the trips.

For each trip at the standard city rate (i.e., within the city limit), the meter computes the fare by combining a base rate of \$2.50, any surcharges, and an incremental charge of \$0.50 for each unit of distance (0.2 miles at a speed of at least 12 miles per hour) or time (60 seconds when the cab is not in motion or is traveling at less than 12 miles per hour).

The common institutional arrangement involves two drivers sharing the same cab. Drivers typically switch shifts at 5 AM and 5 PM, resulting in systematic drops in the number of cabs available in the early morning and early evening (see Appendix Figure 1).[9] The TLC regulates the maximum amount that can be charged to lease a cab for a twelve-hour shift, with a "lease cap" of roughly \$130 depending on the day of the week and the time of the shift.[10]

In addition to institutional constraints, weather can potentially affect labor-supply decisions. Our study uses minute-level weather data (temperature, precipitation, and wind speeds) from the National Centers for Environmental Information collected at five locations around NYC. We match each trip from the TPEP data with the weather conditions at the closest station during the minute when the trip ends.

B. *Descriptive Statistics*

The raw data consist of information on about 41,000 unique drivers and 14,000 taxicabs taking around 173 million trips in 2013. To study cabdrivers' labor-supply decisions, we group trips into *shifts*. We define a shift as a sequence of consecutive trips that are not more than six hours apart from each other (Haggag and Paci, 2014). In other words, we infer that a driver ends a shift after a given trip if the driver does not pick up any more passengers within the next six hours. We define a break as a long waiting time between fares following Farber (2005).[11]

[9]For additional details, see Frechette, Lizzeri and Salz (2016) as well as the January 2011 article by Michael M. Grynbaum in the New York Times: `http://www.nytimes.com/2011/01/12/nyregion/12taxi.html`.

[10]During our sample period, the lease caps for standard vehicles were \$115 for all AM shifts, \$125 for Sunday-Tuesday PM shifts, \$130 for Wednesday PM shifts, and \$139 for Thursday-Saturday PM shifts. The lease caps for hybrid vehicles are \$3 higher. Cabs can also be leased on a weekly basis, with a lease cap that is about six-sevenths of the sum of the daily lease caps.

[11]Specifically, a break consists of a period of at least 30 minutes between a fare that ends in Manhattan and a fare that starts in Manhattan, at least 60 minutes between fares that start or end outside Manhattan but do not end at an airport, or at least 90 minutes between a fare that ends at an airport and the next

Our analysis of daily income effects focuses on cumulative earnings, which we define at the trip level as the sum of fare earnings (excluding tips) from the beginning of the shift to the end of the current trip.[12] After eliminating shifts with missing or inconsistent information (see Appendix A), over 5.8 million shifts by over 37,000 drivers remain, comprising over $1.5 billion in transactions for cab fares.

Summary statistics at the trip level and at the shift level appear in Appendix Table 1. A typical shift consists of 22 trips with a median fare of $9.50. Over 85 percent of all trips start and end in Manhattan, and the average ride takes about 12 minutes. A driver spends over half of the time in a typical shift riding with passengers, 30 percent of the time searching for the next passenger, and 16 percent of the time on break. The majority of shifts last between 7 and 10 hours. Figure 1 displays the fraction of shifts starting at each hour of the day as well as the distribution of work hours.

The market wage varies considerably throughout the day. For each trip, we define the driver's per-minute wage as the ratio of the fare they earn to the number of minutes spent searching for or riding with passengers for that trip. We define the market wage in each minute as the average of the per-minute wages of all drivers working during that minute, plotted in Appendix Figures 2 and 3. Cabdrivers earn a market wage of about $31 per hour, which amounts to a gross income (excluding tips) of about $270 per shift, from which drivers may pay leasing fees and gasoline costs. The highest wages occur during the hours with the lowest number of drivers working, which correspond to the transitions between AM and PM shifts described earlier.

Given our emphasis on the effects of accumulated earnings in the empirical analysis that follows, whether realized wages convey information about future wages can affect the interpretation of our results. We investigate the predictability of hourly wages by residualizing hourly wages on a set of time effects (an interaction between the hour of day and day of week, the week of the year, and an indicator for federal holidays) and weather effects. We find a positive autocorrelation: higher recent earnings predict greater opportunities from continuing (see Appendix Figure 7), which tends to strengthen any evidence that drivers stop working in response to higher accumulated earnings.

II. Tests of Income Effects

A. Stopping Model

We begin by examining the marginal effect of accumulated fare earnings on the probability of ending a shift. We model the decision of a driver at the end of each trip to stop working or to continue working, starting with the specification from Farber (2005). After completing t trips and accumulating y_{int} in fares after

fare.

[12]See Haggag and Paci (2014) and Thakral and Tô (2019) for analyses of tipping decisions.

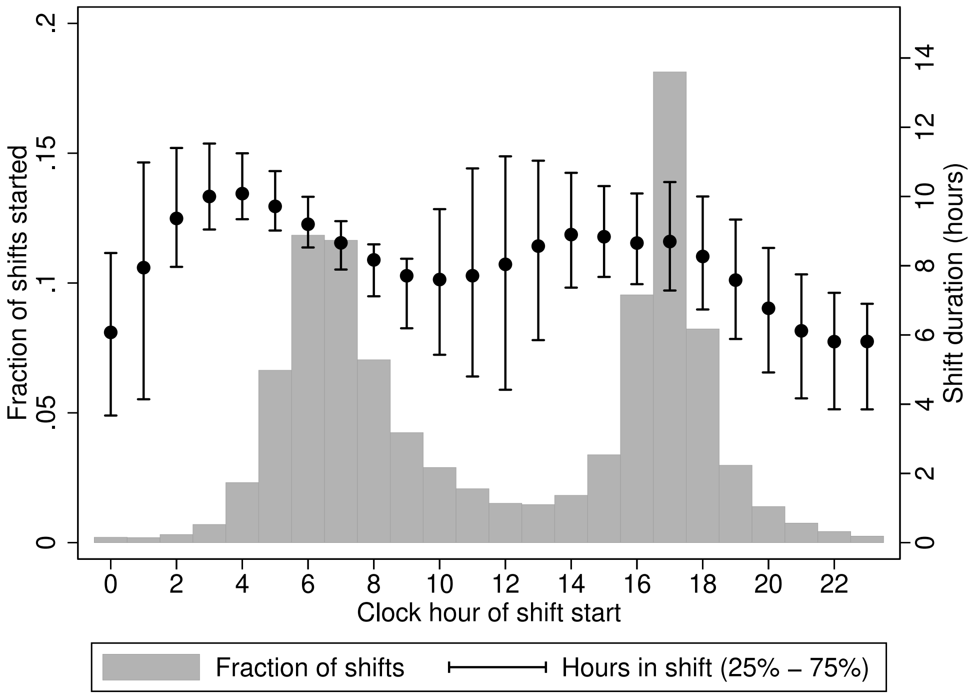

Figure 1. Shift-level summary statistics

Note: The histogram depicts the distribution of shifts by the clock hour of when the shift starts between hour 0 and hour 23. For each clock hour, the distribution of duration of shifts starting at that hour is depicted by the bar graph, with the mean and interquartile range.

a total of h_{int} hours, driver i decides to end shift n when the cost of additional effort exceeds the expected continuation value. Letting d_{int} indicate the decision to stop working, we write the probability that driver i ends shift t at trip n as

$$\text{(F-1)} \qquad \Pr\left(d_{int}=1\right)=\alpha h_{int}+\beta y_{int}+X_{int}\gamma+\mu_i+\epsilon_{int},$$

where X consists of controls that can potentially be related to variation in earnings opportunities from continuing to work, such as location, time, and weather, and μ_i denotes driver fixed effects. The parameter β represents the effect of accumulated earnings on quitting. This discrete-choice problem represents a reduced form of a forward-looking dynamic optimization model based on hours worked so far on the shift, expectations about future earnings possibilities, and other variables that could affect preferences for work.[13]

[13]Buchholz, Shum and Xu (2018) provide a method for estimating a dynamic optimal stopping model, applied to the labor-supply decisions of cabdrivers.

The functional-form assumptions in Equation (F-1) give rise to a potential concern about its ability to deliver consistent estimates of the effect of earnings on quitting. Intuitively, due to the positive correlation between accumulated income and hours of work, a misspecified functional form for the relationship between hours and the stopping probability can cause the model to incorrectly attribute part of the effect of hours to earnings. This issue can also arise when requiring that the marginal effect of additional earnings does not vary with hours worked. For comparison, we reproduce specifications from more recent work:

$$\text{(F-2)} \qquad \Pr\left(d_{int}=1\right)=\sum_j \alpha_j \mathbf{1}_{\{h_{int}\in H_j\}}+\sum_k \beta_k \mathbf{1}_{\{y_{int}\in Y_k\}}+X_{int}\gamma+\mu_i+\epsilon_{int}$$

$$\text{(F-3)} \qquad \Pr\left(d_{int}=1\right)=\sum_{j,k} \delta_{jk} \mathbf{1}_{\{h_{int}\in H_j\}}\mathbf{1}_{\{y_{int}\in Y_k\}}+X_{int}\gamma+\mu_i+\epsilon_{int},$$

where H_j and Y_k form a partition of hours and income, respectively.[14] Farber (2005) estimates Equations (F-1) and (F-2) using a probit model, and Farber (2015) estimates Equations (F-2) and (F-3) using a linear probability model.[15] While these models relax the linearity that Equation (F-1) imposes, adding fixed effects based on a coarse partition of hours may not fully resolve the problem, and misspecification may also persist since the marginal effect of earnings still does not depend on hours worked. When considering neoclassical behavior such as quitting after reaching a fixed number of hours irrespective of income, an empirical Monte Carlo exercise demonstrates that all three of these specifications can yield spuriously significant estimates of β, either positive or negative, as documented in Appendix C.2.

To address the concern arising from restrictive functional forms, we introduce a specification that allows for a flexible, driver-specific hazard of stopping as well as a time-dependent relationship between each of the covariates and the stopping probability:

$$\text{(TT)} \qquad \Pr\left(d_{int}=1\right)=f(h_{int})+\beta(h_{int})y_{int}+X_{int}\gamma(h_{int})+\mu_i(h_{int})+\epsilon_{int},$$

where $f(\cdot)$ represents the baseline hazard and μ absorbs differences in drivers' baseline stopping tendencies.[16] While Equations (F-1) to (F-3) impose that for any pair of drivers one of them has a uniformly higher or lower predicted probability of stopping at the end of any given trip conditional on the other covariates, Equation (TT) accommodates a driver-specific relationship between hours and the probability of stopping. Similarly, Equations (F-1) to (F-3) may suggest that

[14] As in Farber (2015), we take H to partition the shift at 3, 6, 7, 8, 9, 10, 11, 12, and 13 hours and Y at 100, 150, 200, 225, 250, 275, 300, 350, and 400 dollars.

[15] The average marginal effects from the probit model do not materially differ from the estimates of the linear probability model.

[16] Equation (TT) has the form of a non-parametric additive hazards model (Aalen, 1989). See Martinussen and Scheike (2006) for a textbook treatment. Simulation results appear in Appendix C.2.

drivers are more likely to stop at 4 PM, when it rains, or when a trip ends near the taxi garage regardless of how many hours they have worked, whereas Equation (TT) allows the marginal effect of each variable on the probability of stopping to vary flexibly throughout the shift. The flexibility of Equation (TT) comes at the cost of requiring more data.

The term $\beta(h)$ represents the effect of an additional dollar of accumulated daily earnings on the probability of ending a shift for a driver who finishes a trip after h hours of work. A positive effect of accumulated earnings on quitting suggests the presence of a daily income effect under the assumption that cumulative daily earnings are uncorrelated with unobserved determinants of the value of stopping (such as effort or fatigue) or the value of continuing (such as future earnings opportunities) conditional on the full set of time-varying covariates, which Section II.D discusses in more detail.

We use local linear regression techniques to estimate the baseline hazard and the time-varying coefficients in Equation (TT). For any given time h, the associated parameter estimates solve a separate weighted least squares problem (Cleveland and Devlin, 1988)

$$\min_{\alpha,\beta,\gamma,\mu_i} \sum_{i,n,t} w(h_{int} - h) \left(d_{int} - (\alpha h_{int} + \beta y_{int} + X_{int}\gamma + \mu_i)\right)^2$$

with weights given by the function $w(\cdot)$. Using uniform weights, the coefficients at any time h represent the fit of a linear model to a localized subset of the data. The results we report in Section II.B use uniform weights over a 10-minute window of time during the shift.[17] This approach uses only variation due to trips ending within 10 minutes of a given h to provide ordinary least squares estimates of the corresponding parameters in Equation (TT), while Equations (F-1) to (F-3) make use of the full set of trips for estimating a more limited set of parameters.

B. *Estimates of the Stopping Model*

Table 1 presents in Panel A our estimates of the elasticity of stopping at 8.5 hours (approximately the median stopping time) with respect to accumulated fare earnings. The strategy for estimating the quitting response to additional accumulated earnings follows Farber (2005) in using variation in earnings conditional on an extensive set of covariates that capture the value of stopping (hours worked so far on the shift) and the value of continuing (expectations about future earnings possibilities). Appendix D.1 provides more detail on variation in earnings, and Section II.D discusses a supplementary analysis that uses speeds to instrument for earnings.

The first column reports estimates from our preferred specification, based on Equation (TT). Each row corresponds to a more comprehensive set of controls

[17] We find that varying the window (e.g., to 5 or 30 minutes) or using a local quadratic fit results in similar estimates for $\beta(h)$.

Table 1—Elasticity of stopping at 8.5 hours with respect to income

Panel A: Effect of 1% increase in accumulated earnings	(TT)	(F-1)	(F-2)	(F-3)
Controlling for				
Hours	0.0636	0.7853	0.0041	-0.2385
	(0.0272)	(0.0143)	(0.0495)	(0.0931)
& Drivers	0.4838	0.8748	0.0754	-0.1999
	(0.0187)	(0.0123)	(0.0486)	(0.0915)
& Time	0.8304	0.6689	0.4902	0.2494
	(0.0176)	(0.0132)	(0.0468)	(0.0879)
& Location	0.3355	0.2078	0.3764	0.1186
	(0.0169)	(0.0132)	(0.0461)	(0.0865)
& Weather	0.3336	0.2074	0.3759	0.1183
	(0.0169)	(0.0132)	(0.0461)	(0.0865)

Panel B: Comparison with previous estimates (95% C.I.)

This paper:	0.3336	(0.3005, 0.3667)
Farber 2005:	0.1210	(-0.2023, 0.4443)
Farber 2015:	0.9066 (day), 0.0689 (night)	

Note: Panel A reports in each cell an estimate of the percent change in the probability of ending a shift at 8.5 hours in response to a 1 percent increase in cumulative earnings. The columns corresponds to the specifications in Equation (TT) and Equations (F-1) to (F-3), respectively. All specifications include controls for minutes spent working, including indicators for whether the driver spends time with passengers in each hour. Time controls include fixed effects for hour of day by day of week and for day of year. Location controls consist of neighborhood fixed effects. Weather controls consist of indicators for precipitation, wind speed, and temperature in the minute that a trip ends. Drivers denotes fixed effects for the anonymized license numbers. Standard errors reported in parentheses are adjusted for clustering at the driver level. Panel B reports calculations based on Table 5 of Farber (2005) and Table VII of Farber (2015) as explained in Footnotes 20 and 21. The sample consists of over 37,000 drivers, with all rows of the first column reflecting estimates from trips that end within 10 minutes of 8.5 hours (2.3 million trips) and all rows of the remaining columns reflecting estimates from a two-fifteenths sample of all trips (16 million trips). See Appendix A for further details.

than the previous one. All specifications consist of controls for minutes spent working, including indicators for whether the driver spends time with passengers in each hour. The specification in the first row, with no additional controls, shows a small but statistically significant relationship between cumulative daily earnings and stopping probabilities. If drivers with higher average earnings tend to work more, then using across-driver variation likely underestimates the relationship between accumulated earnings and quitting. The second row shows that accounting for driver-specific stopping tendencies strengthens the estimated effect considerably. Relatedly, to the extent that drivers work more on days with higher expected wages, failing to control for time effects may also understate the magnitude of the income effect. Adding an interaction between clock hour and day of week as well as indicators for day of year in the third row indeed results in a larger estimate. Since drivers may end their shifts with higher probability when a trip ends in a convenient location coinciding with higher accumulated earnings (e.g., near the driver's home or the cab garage in one of the outer boroughs), the fourth row adds fixed effects for the 195 Neighborhood Tabulation Areas (NTA) in NYC (see Haggag, McManus and Paci, 2017), which decreases the estimated effect to a 3.1 percent increase in the probability of ending a shift at 8.5 hours (0.44 percentage-point increase relative to a baseline stopping probability of 13.2 percent) in response to a 10 percent ($26) increase in cumulative earnings.[18] The elasticity estimate remains stable around 0.3 after adding a set of weather controls (indicators for precipitation, temperature above 80 degrees Fahrenheit, temperature below 30 degrees Fahrenheit, and wind speed on the Beaufort scale), measured in the minute when a trip ends, in the last row.

The remaining columns in Panel A report estimates based on Equations (F-1) to (F-3).[19] While the control variables tend to influence the estimates in the directions discussed above, the magnitudes differ across specifications. Since Farber (2015) does not use location or weather controls in estimating the stopping model, the third row of the third column corresponds to the primary specification that Farber (2015) uses for counterfactual analysis and reports an effect that exceeds our preferred estimate by over 40 percent. The less constrained specification in Farber (2015) corresponds to the third row of the fourth column, which reports a smaller but imprecisely estimated effect that does not significantly differ from our preferred estimate. Under the full set of controls, the point estimates across specifications suggest that the probability of ending a shift at 8.5 hours increases by between 1.2 and 3.8 percent in response to a 10 percent increase in cumulative earnings.

Panel B provides a direct comparison with previous papers. Farber (2005) reports a statistically insignificant effect of earnings on quitting (reproduced in Panel B), though the point estimate implies that a 10 percent increase in cumulative earnings corresponds to a 1.2 percent increase in the probability of ending a shift.[20] The confidence interval also encompasses the estimates from all specifications with the full set of controls in Panel A, including our preferred estimate. Farber (2015), using a specification analogous to Equation (F-2), reports a separate estimate for day shifts (start between 4 AM and 10 AM) and night shifts (start between 2 PM and 8 PM) and finds sizable income effects only for day shifts (9 percent increase in the probability of stopping at 8.5 hours in response

[18] Allowing for driver-specific location effects strengthens the result (see Appendix Table 7 column 1).

[19] Since Equations (F-2) and (F-3) coarsen income into bins of at least $25, we compute the effect of a 10 percent increase in income on the probability of stopping after earning $260 in 8.5 hours and scale this to obtain the elasticity.

[20] Farber (2005) reports that an additional $100 increases the probability of ending a shift by 0.011 (s.e. 0.015) at 8.5 hours under the full set of controls (Table 5). With a mean income of $161.33 (Table B1), an additional 10 percent in earnings corresponds to a $16.13 \cdot 0.011 \approx 0.18$ (s.e. 0.24) percentage-point increase in the probability of stopping relative to a baseline of 14.67 percent (Table 4), or a 1.210 (s.e. 1.650) percent increase in the probability of ending a shift.

to 10 percent higher cumulative earnings).[21] Appendix Table 5 replicates Table 1 for day shifts and night shifts and confirms this pattern for the Farber (2015) specifications (columns 3 and 4), finds the opposite pattern for the main Farber (2005) specification (column 2), and shows that our preferred approach (column 1) instead yields very similar estimates for day shifts and night shifts. Overall, Equation (TT) provides evidence of a modest-sized daily income effect, comparable with but more consistent than previous estimates.

Figure 2 shows that the magnitude of the income effect from Table 1, evaluated at 8.5 hours of work, persists throughout the shift. The figure plots the probability of stopping (left axis) and estimates based on Equation (TT) of the percentage-point change in the stopping probability in response to a 10 percent increase in earnings (right axis) every fifteen minutes over a six-hour period, roughly corresponding to the 10^{th} and 90^{th} percentile of the distribution of stopping times. A clear, stark relationship appears between hours of work and the probability of ending a shift, consistent with the prediction of the neoclassical model as Farber (2015) highlights. In addition, the elasticity of stopping with respect to earnings remains significant and consistent throughout the shift as the average stopping probability increases from 2 percent to 28 percent. The figure shows an elasticity of about one-third during the hours leading up to the median stopping time and about one-fourth in the hours that follow.

C. The Role of Timing

We proceed to test whether drivers exhibit stronger responses to more recent experiences. Relaxing the implicit assumption that money is fungible within the day, we augment Equation (TT) to express the probability of stopping as

$$\Pr\left(d_{int}=1\right) = f_j(h_{int}) + \sum_{\ell} \beta^{\ell}(h_{int}) y_{int}^{\ell} + X_{int}\gamma(h_{int}) + \mu_i(h_{int}) + \epsilon_{int}, \tag{1}$$

where y_{int}^{ℓ} denotes fare earnings accumulated in hour ℓ of the shift. If drivers compare their cumulative daily earnings with a fixed target, then we would expect to find that the impact of an additional dollar on the probability of ending a shift does not depend on when the dollar was accumulated (i.e., β^{ℓ} is independent of ℓ).

Table 2 presents estimates of the percent change in the probability of ending a

[21] According to Table VII in Farber (2015), a \$25 increase in income during a day shift from \$225–\$249 to \$250–\$274 increases the probability of ending a shift by $0.0389 - 0.0264 \approx 0.0125$ (standard error for the difference not reported). With a mean income of \$248.41 (Table III), an additional 10 percent in earnings corresponds to a $24.84 \cdot 1.25/25 \approx 1.242$ percentage-point increase in the probability of stopping relative to a baseline of 13.7 percent, or a 9.066 percent increase in the probability of ending a day shift. For night shifts, a \$25 increase in income from \$250–\$274 to \$275–\$299 increases the probability of ending a shift by $-0.0033 - (-0.0042) \approx 0.0009$. With a mean income of \$262.03 (Table III), an additional 10 percent in earnings corresponds to a $26.20 \cdot 0.09/25 \approx 0.0943$ percentage-point increase in the probability of stopping relative to a baseline of 13.7 percent, or a 0.689 percent increase in the probability of ending a night shift.

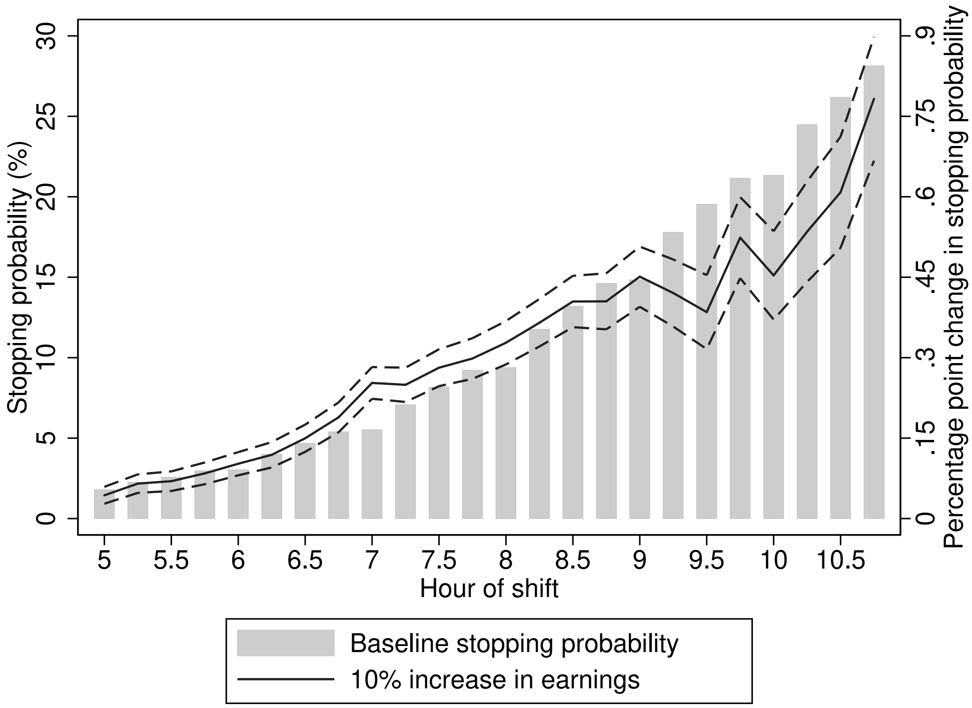

Figure 2. Stopping model estimates: Income effect throughout the shift

Note: The bars, corresponding to the scale on the left, show the probability that a driver ends a shift after completing a trip at the specified number of hours. The solid lines, corresponding to the scale on the right, depict the percentage-point change in the probability of stopping, evaluated at various times throughout the shift, in response to a 10 percent increase in earnings. Estimates obtain from Equation (TT) with the full set of controls (see Table 1 for details) and fixed effects for over 37,000 drivers. Each point represents estimates from trips that end within 10 minutes of the specified number of hours on the horizontal axis (ranging from 3.5 million trips to 0.7 million trips). The dashed lines represent the 95-percent confidence interval, with standard errors adjusted for clustering at the driver level.

shift at 8.5 hours in response to an additional 10 percent ($26) in income earned at various times in the shift based on Equation (1). While a 10 percent increase in cumulative earnings corresponds to a 3 percent increase in the probability of stopping overall, we see that earnings accumulated in the first few hours of the shift do not increase the probability of ending a shift at 8.5 hours. If the additional earnings arrive in the eighth hour of the shift, then our estimates imply a 10.2 percent increase in the probability of stopping (1.35 percentage-point increase relative to a baseline stopping probability of 13.2 percent) under the full set of control variables.[22] Figure 3 summarizes the main result, showing how

[22]Extending Equations (F-1) to (F-3) to include an effect of earnings in each hour also results in the strongest effects for the most recent earnings, as Appendix Table 8 shows.

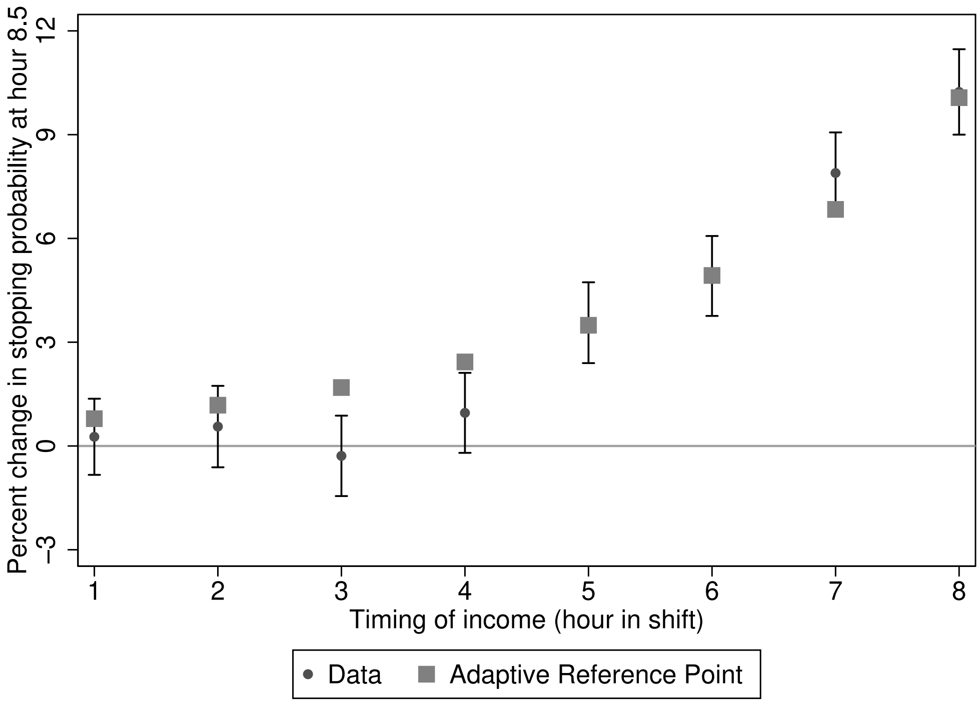

Figure 3. Stopping model estimates: Elasticity of stopping at 8.5 hours with respect to income—By timing of income

Note: The figure depicts the percent change in the probability of ending a shift at 8.5 hours (baseline stopping probability of 13.2 percent) in response to a 10 percent ($26) increase in earnings accumulated at different times in the shift. The circles and lines represent point estimates and 95 percent confidence intervals from Equation (1) with the full set of controls and fixed effects for over 37,000 drivers, as reported in Table 2 column (5). The gray squares represent predictions from the estimated model of adaptive reference points (see Table 3).

the effect gradually diminishes for earnings accumulated earlier.[23] An additional dollar accumulated in the eighth hour increases the probability of stopping by an order of magnitude more than an additional dollar accumulated four hours earlier. For comparison, an additional 10 minutes of work (median trip duration) increases the probability of ending a shift by over 7 percent.

Figure 4 shows that the recency pattern from Figure 3, evaluated at 8.5 hours of work, persists throughout the shift. The columns of the figure correspond to different times during the shift between 5.5 and 11 hours. Each column depicts the effect that an additional 10 percent of cumulative income, accumulated in different hours corresponding to the rows, has on the probability of stopping.

[23]The gray squares in Figure 3 display the prediction from the model, which we discuss in Section III.

TABLE 2—ELASTICITY OF STOPPING AT 8.5 HOURS WITH RESPECT TO INCOME—BY TIMING OF INCOME

	(1)	(2)	(3)	(4)	(5)
Income in hour 1	-0.7436	-0.8116	0.0658	0.0305	0.0266
	(0.0884)	(0.0624)	(0.0576)	(0.0561)	(0.0561)
Income in hour 2	-1.8319	-0.5435	0.0504	0.0593	0.0561
	(0.0780)	(0.0658)	(0.0615)	(0.0600)	(0.0601)
Income in hour 3	-0.6928	0.0088	-0.0302	-0.0254	-0.0286
	(0.0731)	(0.0648)	(0.0609)	(0.0593)	(0.0593)
Income in hour 4	-0.0717	0.5183	0.1255	0.0993	0.0957
	(0.0720)	(0.0643)	(0.0607)	(0.0590)	(0.0590)
Income in hour 5	0.3788	0.9209	0.3828	0.3589	0.3564
	(0.0734)	(0.0649)	(0.0612)	(0.0596)	(0.0596)
Income in hour 6	0.0129	0.4357	0.5215	0.4939	0.4915
	(0.0736)	(0.0646)	(0.0604)	(0.0589)	(0.0589)
Income in hour 7	-0.0339	0.1998	0.8238	0.7908	0.7891
	(0.0739)	(0.0658)	(0.0614)	(0.0598)	(0.0598)
Income in hour 8	0.7454	0.7121	1.651	1.0242	1.0235
	(0.0763)	(0.0705)	(0.0651)	(0.0630)	(0.0630)
Controls					
Hours	X	X	X	X	X
Drivers		X	X	X	X
Time			X	X	X
Location				X	X
Weather					X

Note: The table reports estimates from Equation (1) of the percent change in the probability of ending a shift at 8.5 hours in response to a 1 percent increase in accumulated daily earnings that arrive in the specified hour. Each column adds an additional set of control variables, described in Table 1. Standard errors reported in parentheses are adjusted for clustering at the driver level. The estimates represent data from trips that end within 10 minutes of 8.5 hours (2.3 million trips) from over 37,000 drivers.

Shifts that start at different times of the day exhibit similar qualitative effects, as Appendix Figure 6 shows. The recency pattern also holds for groups of drivers with lower or higher variability of daily hours, with especially pronounced effects for the latter group (see Appendix Table 6) which likely faces fewer constraints related to stopping times. The results are robust to controlling for driver-specific location effects (see Appendix Table 7 column 1). In addition to finding that stopping decisions do not respond to earnings accumulated early in a shift, we also see no significant relationship between the probability of ending a shift and earnings on the previous day (see Appendix Table 7 column 2).

D. *Alternative Explanations*

This section addresses potential challenges to our interpretation of the excess sensitivity of daily labor-supply decisions to recent earnings as evidence of a time-varying income effect. We consider the possibility that cumulative earnings are correlated with unobserved determinants of the stopping decision such as effort and fatigue, or that cumulative earnings convey information about future earnings opportunities. In addition, we assess whether the relationship between recent earnings and quitting arises due to other factors such as driver inexperience, liquidity constraints, and mismeasurement of work hours.

Effort and fatigue

A potential concern with our interpretation would arise if additional recent earnings coincide with an increase in the intensity or difficulty of work. We first discuss how our estimates suggest limited scope for the effect of recent earnings to reflect a response to fatigue, and we then apply an instrumental-variable (IV) approach to address the concern more directly.

If fatigue poses a confound for estimating the effect of recent earnings, we would expect to find much larger magnitudes in the later hours of a shift (e.g., after working 10 hours compared to 8.5 hours) insofar as drivers face an increasing marginal disutility of effort, and we would expect to find much weaker patterns after adding a flexible, driver-specific relationship between work hours and the probability of stopping. The fact we find consistent and sizable recency effects *throughout* the shift (Figure 4) presents a difficulty for fatigue-based explanations. Moreover, the consistency of results *across* specifications (Table 2 and Appendix Table 8) further suggests that the estimated effect of recent earnings does not reflect a response to fatigue.

To address directly the possible correlation between recent earnings and unobserved determinants of the decision to end a shift, we propose an IV approach based on a proxy for traffic conditions. Better traffic conditions enable drivers to cover greater distances in a given amount of work time and thus accumulate higher earnings. However, drivers may also accumulate higher earnings due to increases in effort, which directly contributes to their quitting decisions. To capture variation in earnings at different times due to traffic conditions plausibly unrelated to a driver's decision to exert additional effort, we instrument for earnings in each hour based on contemporaneous speeds of nearby drivers. Although the IV estimates are noisier, we continue to find strong effects for the most recently accumulated earnings (see Appendix D.1 for more detail).

Learning about Future Earnings

Another potential concern with interpreting the effect of earnings on stopping behavior would arise if accumulated earnings convey additional information about future opportunities, either within the same shift or across shifts.

If higher recent earnings signify a lower continuation value conditional on all the covariates, then the estimated relationship between earnings and quitting would overstate the true income effect. However, we find that higher-than-expected recent earnings should, if anything, be associated with a higher value of continuing to work because residualized earnings exhibit a positive autocorrelation (Appendix Figure 7).

Likewise if higher earnings correlate with plentiful opportunities on the next day, then drivers may engage in intertemporal substitution, quitting during times of high earnings to conserve energy for the next shift. However, we find that higher-than-expected earnings do not appear predictive of market conditions on subsequent days (Appendix Figure 8).

Driver Heterogeneity: Experience and Liquidity Constraints

We investigate whether the positive relationship between earnings and stopping merely reflects a failure to optimize by inexperienced drivers.[24] Given that performance improves quickly with experience in this setting (Haggag, McManus and Paci, 2017), drivers might also learn to supply labor more efficiently by ignoring daily earnings. Classifying drivers as *new* if we first observe them in our data on or after April 1 (Haggag, McManus and Paci, 2017), we do not find any evidence of larger quitting responses to recent earnings for new drivers (see Appendix Table 10). In addition to this across-driver definition of experience, we also consider a within-driver definition of experience and do not find significant differences as new drivers gain more experience (see Appendix Table 11).

We also investigate whether liquidity constraints explain the patterns we observe by replicating our analysis on a sample of drivers for whom liquidity constraints likely do not bind. Specifically, we estimate the stopping model restricted to owner-drivers, as such drivers possess enough borrowing power or wealth to purchase an independent medallion to operate a taxicab. The estimates suggest that liquidity constraints do not confound the income effects we observe (Appendix Table 7 column 3). In fact, owner-drivers exhibit stronger income effects, consistent with our finding of stronger income effects for drivers with more flexibility in their hours decisions (Appendix Table 6).

Measurement of Work Hours

Two issues arise when measuring work hours in this setting. First, the data do not distinguish between a driver who ends a shift immediately after dropping off their last passenger and a driver who spends time searching for another fare unsuccessfully. This would only pose a concern if drivers tend to face greater difficulties in finding passengers towards the end of shifts in which they earn more. Second, the data do not contain an explicit measure of break times. A

[24]Camerer et al. (1997) propose the idea that more experienced drivers might exhibit more positive wage elasticities of labor supply, and the evidence from Farber (2015) corroborates this.

positive correlation between recent earnings and unobserved tiredness, and thus quitting behavior, could arise if the additional earnings induce drivers to take fewer breaks. However, the data show the opposite of this pattern. Appendix D.5 discusses both of these issues in more detail.

III. Structural Model of Reference Dependence

Having established the existence of a recent-earnings effect, we develop and estimate a model of adaptive reference points that can explain this effect. While various different models can potentially formalize how reference points influence behavior, this section focuses on models based on prospect theory, as existing work invokes reference dependence and loss aversion to explain income-targeting behavior. Appendix F provides a complementary investigation of how models based on salience (Bordalo, Gennaioli and Shleifer, 2015) account for the evidence. Even without taking a strong stance on a particular account of how the reference level influences decisions, our findings suggest that the reference level must adjust within a day. We therefore aim to develop and assess alternative formulations of the reference point.

A. Daily Labor Supply Model Setup

A General Stopping Model

At the end of trip t, a driver with cumulative earnings I_t and hours of work H_t chooses whether to stop or continue working (Farber, 2008). The driver decides to quit for the day if the value of stopping exceeds the expected value of continuing to work, i.e.,

$$\mathbb{E}_t\left[v(I_{t+1}, H_{t+1})\right] - v(I_t, H_t) + \varepsilon_t < 0, \tag{2}$$

where ε is an error term. The expected value of continuing depends on the joint distribution of the fare f_{t+1} and duration h_{t+1} of trip $t+1$, given the circumstances at the end of trip t. The error term comprises a vector of variables that determine the difference between the current utility and the continuation value as well as an idiosyncratic component.

A Neoclassical Model

As a benchmark, we consider the case that the marginal utility of lifetime income does not vary in response to small, within-day changes in wealth (see Appendix C.1). To capture this, the objective function v takes the form

$$\begin{aligned} v(I_t, H_t) &= v_I(I_t) + v_H(H_t) \\ &= I_t - \frac{\psi}{1+\nu} H_t^{1+\nu}, \end{aligned} \tag{3}$$

where ψ parameterizes the disutility of work and ν is the elasticity parameter. If drivers make decisions according to a quasi-linear objective function, then labor supply does not decrease in response to additional accumulated earnings.[25]

A Reference-Dependent Model

Explaining the results in Section II requires non-trivial within-day changes in the marginal utility of income. A back-of-the-envelope calculation suggests that the daily income effect we observe implies an implausibly high degree of risk aversion of over 100 (see Appendix E.1). The leading explanation in the literature for the mixed evidence on behavior in daily labor-supply decisions comes from the Kőszegi and Rabin (2006) theory of reference-dependent preferences (see Crawford and Meng 2011 and a survey of the earlier work in DellaVigna 2009). To build on this, we proceed by describing an objective function v^{LA} for a given reference point, and we address different models of the reference point in Section III.B.

In the model, utility depends not only on a standard outcome-based consumption component but also on a gain-loss component which captures how decision makers assess choices relative to a reference point. The objective function of the driver takes the form

$$v^{\text{LA}}(I_t, H_t) = (1-\eta)v(I_t, H_t) + \eta n\left(I_t \,|\, I_t^r\right), \tag{4}$$

where I^r denotes the reference level for income (i.e., the driver's expected earnings for the shift), η determines the relative weight on gain-loss utility, and the gain-loss utility is given by

$$n\left(I \,|\, I^r\right) = \left(\mathbf{1}_{\{I>I^r\}} + \lambda \mathbf{1}_{\{I<I^r\}}\right)\left(I - I^r\right),$$

where $\lambda \geq 1$ parameterizes the degree of loss aversion. This coincides with the neoclassical model when there is no difference in utility from gains and losses (i.e., $\lambda = 1$ or $\eta = 0$). Compared to the specification from Crawford and Meng (2011), Equation (4) does not include a reference level for hours of work. In addition, this formulation makes two simplifying assumptions about the general gain-loss component of utility from Kőszegi and Rabin (2006). First, the piecewise-linear gain-loss function rules out diminishing sensitivity, the observation that decision makers experience smaller marginal changes in gain-loss sensations further away from their reference levels. Second, the reference levels represent a driver's point expectations for income and hours on a given shift, abstracting from stochasticity whereby the reference levels represent the full distribution of potential earnings for that particular shift. The results we present do not substantively change when we relax each of these assumptions in Section III.D.

[25] This conclusion holds as long as the continuation value does not decrease in I_t. See the positive autocorrelation in Appendix Figure 7 for empirical evidence supporting this.

B. Specifications of the Reference Point

Loss aversion produces a daily income effect, exhibited by a reduction in labor supply, through a decrease in the marginal utility of income at a given reference point. The specification of the reference point crucially affects predictions about how labor supply responds to the timing of earnings. Note that an exogenous reference point (e.g., a round number such as \$250, or twice the daily fee for leasing the cab, as Camerer et al. 1997 informally propose) predicts no role for the timing of earnings, in stark contrast with our empirical results. We thus proceed by characterizing the dynamics of the reference point.

One class of specifications consists of forward-looking reference points based on the lagged expectation. Kőszegi and Rabin (2006) posit that rational expectations endogenously determine the reference point. They make an important distinction between how beliefs and preferences adjust to new information: the model does not require expectations to adjust slowly, but reference points depend on the lagged expectation and thus do not change instantaneously with new information.

In the absence of a theoretical account of how quickly the reference point adjusts, we consider a range of possibilities. A lag sufficiently long that reference points do not adjust results in a fixed reference point as in the original view of income targeting (Camerer et al., 1997). With a vanishingly short lag, reference points fully adjust to new information, removing the influence of the reference point on decisions and resulting in neoclassical behavior. In general, we take the reference point in period τ as the expectation held in period $\tau - 1$, for some definition of a period. The daily-level income target from (Crawford and Meng, 2011), which we denote I_0^r, lies between the two extremes above, as does a reference point that updates every hour or every trip.

Another class of specifications consists of backward-looking reference points based on past experiences or outcomes.[26] Post et al. (2008) presents a dynamic model for the reference point which separates the effect of initial expectations from the effect of the most recent outcomes. In our setting, a reference point along these lines preserves some of the advantages of a forward-looking reference point, such as explaining why reference dependence does not require that higher wages generically lead to lower effort.

We model the adaptive reference point as a convex combination of the lagged reference point and the reference point that would obtain if new information were fully incorporated. This requires a way of defining new information at trip t, which we denote Δ_t, as well as the expectation of daily earnings after trip t, which we denote E_t. The quantity Δ_t represents the difference between realized earnings I_t and expected earnings. We take a simple approach to computing expected earnings by predicting the fare f_t and ride duration h_t based on conditions at the end of trip $t - 1$ (time, location, and weather), and we adjust the expected fare

[26]For instance, Bowman, Minehart and Rabin (1999) and DellaVigna et al. (2017) consider reference points based on past consumption or income.

based on the realized duration of trip t so that $\Delta_t = f_t - \mathbb{E}_{t-1}[f_t]\frac{h_t}{\mathbb{E}_{t-1}[h_t]}$.[27] As a reduced form for the updating problem that characterizes expectations of daily earnings, we write $E_t = I_0^r + \sum_{\tau=1}^{t} \Delta_\tau$.

We define the updated reference point as

$$I_t^r = \theta I_{t-1}^r + (1-\theta)E_t, \tag{5}$$

where $0 \leq \theta \leq 1$, with $\theta = 1$ corresponding to a reference point that does not adjust within the day and $\theta = 0$ corresponding to a reference point that adjusts instantaneously. Rewriting this recursive formulation, we express the updated reference point as

$$I_t^r = \theta^t I_0^r + (1-\theta)\sum_{\tau=1}^{t} \theta^{t-\tau} E_\tau,$$

which highlights that the adaptive reference point consists of a weighted average of multiple lagged values of expectations. Rewriting the updated reference point as

$$I_t^r = I_0^r + \sum_{\tau=1}^{t}(1-\theta^{t+1-\tau})\Delta_\tau.$$

highlights that the reference point incorporates less recent earnings to a greater extent, consistent with the idea that reference points take time to adjust in response to recent changes in expectations.

Explaining the recency effect requires a slow-adjusting reference point within the day.[28] While our specification nests the static reference point ($\theta = 1$) as well as a reference point that adjusts instantaneously ($\theta = 0$), both of these extreme cases imply the fungibility of money within a shift, and only the intermediate case with $0 < \theta < 1$ permits a violation of fungibility. A reference point that does not evolve within the day can account for a labor-supply response to earnings but eliminates the scope for more recent experiences to have a stronger influence on stopping decisions. With a reference point that adjusts within a shift, a more recent gain may make a driver more likely to exceed his income target than an earlier gain because the reference point takes time to adjust. However, a reference point that adjusts instantaneously cannot explain the sensitivity of labor-supply decisions to daily earnings because deviations from expectations no longer bring cumulative daily earnings closer to or further from the reference point.

[27] As an alternative approach, we could add the realized duration of trip t as a predictor of f_t (i.e., set Δ_t to $\mathbb{E}_{t-1}[f_t \mid h_t]$), and we could also add lagged income and work hours as predictors of f_t.

[28] For reduced-form evidence that the probability of stopping significantly increases when income exceeds a target that updates more in response to earlier experiences, pointing towards an adaptive reference point, see Appendix E.2.

C. Estimation and Identification

Using the data on stopping decisions, we estimate the models via maximum likelihood under the various specifications of the reference point. Given the stopping rule in Equation (2), we follow Crawford and Meng (2011) in assuming that $\varepsilon_t = x_t\beta + \xi_t$, where $x_t\beta$ captures the effect of control variables and ξ_t are independent and normally distributed with mean zero and variance σ^2. This yields likelihood functions of the form

$$\sum \log \Phi\left(\frac{v^{\mathrm{LA}}(I_t, H_t) - \mathbb{E}_t\left[v^{\mathrm{LA}}(I_{t+1}, H_{t+1})\right]}{\sigma}\right), \tag{6}$$

where Φ denotes the standard normal cumulative distribution function. The control variables consist of the time, weather, and location controls from Section II. To accommodate a more flexible relationship between hours and quitting, we allow the disutility of effort to take a separate value on each half-hour interval of the shift.[29]

We begin with the adaptive reference point from Equation (5). To proxy for drivers' initial expectations I_0^r, we follow Crawford and Meng (2011) in using the sample average of income by driver and day of week (excluding the current shift). To set the updating term Δ_t, we define trip-level expectations of fares f_t based on a regression of next-trip fare on the time, location, and weather controls from Table 1 (and similarly for hours h_t).[30] Under the specifications of the reference point that we consider, the parameters η and λ are not separately identifiable because the decision maker takes the reference point as exogenous to their choice, and thus behavior depends only on the ratio between utility from losses and gains $\Lambda = 1 + (\lambda - 1)\eta$ (see Appendix E.3 for more detail).

Although we estimate the parameters jointly, the following describes the key sources of variation in that data for identifying each of the parameters. Variation in work hours, expected wages from continuing, cumulative income, and the timing of income contributes to the identification of the disutility of effort ψ, elasticity parameter ν, coefficient of loss aversion Λ, and speed of adjustment θ. The model provides a strong link between these sources of variation in the data the structural parameters, which we discuss in Appendix E.4 and illustrate in Appendix Figure 10.

We estimate the parameters and obtain standard errors using subsampling (Politis and Romano, 1994), which reduces the computational burden. The estimates in Table 3 obtain from 230 subsamples (without replacement) of 150,000

[29] Although the model imposes the same vector of parameters across shifts for the disutility of effort, the true disutility of effort might vary with expected hours for a given shift. To reduce the need for introducing additional parameters, we restrict the sample by removing shifts in the top and bottom quartile of the distribution of expected hours. We proxy for expectations about hours by using the sample average of hours by driver and day of week, excluding the current shift (Crawford and Meng, 2011).

[30] We use seemingly unrelated regressions following Crawford and Meng (2011).

observations each.

Table 3—Maximum likelihood estimates

	(1)	(2)	(3)
Disutil. of effort ψ	0.1105	0.1932	0.1531
	(0.0697)	(0.1480)	(0.0900)
Elasticity ν	0.8585	0.7902	0.8477
	(0.2641)	(0.2582)	(0.2378)
Loss aversion Λ		1.9992	2.6074
		(0.2364)	(0.2968)
Adaptation θ			0.8227
			(0.0640)
Error term distribution σ	0.2487	0.4077	0.4009
	(0.0289)	(0.0914)	(0.0640)
Mean log likelihood	-27,302	-27,056	-26,847
Test $\Lambda = 1$		< 0.001	< 0.001
Test $\theta = 1$			< 0.001

Note: This table presents maximum likelihood estimates of Equation (6) with the adaptive reference point given by Equation (5). The estimation sample consists of 34.5 million trips from over 37,000 drivers. Column (1) corresponds to the restriction $\Lambda = 1$, column (2) corresponds to the restriction $\theta = 1$, and column (3) presents the full specification. We report the estimated disutility of effort parameter that applies to trips that occur between hour 8.5 and hour 9 of the shift. The mean log likelihood reports the average of the log likelihood across the subsamples. The last two rows contain p-values from likelihood ratio tests of the following null hypotheses: (i) the model without loss aversion, and (ii) the model with a static reference point.

Column (1) reports estimates under the restriction $\Lambda = 1$, corresponding to the case without loss aversion, and Column (2) allows for loss aversion relative to a static reference point (i.e., $\theta = 1$). A likelihood ratio test rejects the null hypothesis of no loss aversion. The static specification resembles the model that Crawford and Meng (2011) estimate (see their Table 4 column 5), and the parameter estimates generally fall within the same range. In particular, we find a similar coefficient of loss aversion over income of about 2. The assumption that the reference point does not adjust within the day may lead to an underestimate of the coefficient of loss aversion given our result that drivers react more strongly to more recent earnings.

Relaxing the assumption of a static reference point, column (3) presents estimates from the full specification with loss aversion relative to the adaptive reference point. The results highlight the importance of within-day reference-point adaptation, suggesting that drivers do not treat earnings accumulated at different times as fungible. The estimate of θ differs significantly from 0 and 1, and a likelihood ratio test rejects the restriction to a static reference point. The estimated coefficient of loss aversion increases as expected, and the remaining parameters do not substantially differ from their counterparts in column (2). The magnitude of our estimate for the speed of adjustment θ depends on the definition of a period. Since we take each period to be a trip, the point estimate implies that the reference point adjusts immediately to incorporate about 18 percent of a shock to earnings. Estimating the speed of reference point adjustment at a lower frequency (e.g., defining a period as an hour instead of a trip) would result in a smaller magnitude of θ. Within an hour, the reference point incorporates about 40 percent of the shock, and within four hours only about 10 percent of the shock remains unincorporated.

To assess how well the model fits the data, Figure 3 compares the observed and predicted recency pattern of the income effect. While the stronger effects of recently accumulated earnings serve as the motivation for developing a model in which the reference point adjusts, our estimation approach does not explicitly target these moments. The gray squares depict the predicted effect of an additional 10 percent ($26) in income earned at various times in the shift on the probability of ending a shift at 8.5 hours, with the estimates from Equation (1) providing a benchmark for comparison. Adaptive reference points, as we expect, lead to stronger labor-supply reductions in response to more recent experiences. Moreover, the consistency between the predicted income effects and the observed magnitudes in the data provides support for the model of adaptive reference points.

D. Variants of the Model

The first part of this section evaluates adaptive reference points in comparison with alternative specifications of the reference point. We consider reference points based on the lagged expectation of earnings, ranging from a fixed reference point to one that updates each trip. The second part of this section shows robustness of the main conclusion of an intermediate degree of adaptation to some of the ancillary assumptions of the model. We relax the simplifying assumptions from Section III.A, allowing for reference dependence over hours, diminishing sensitivity, and stochastic reference points.

Alternative Specifications of the Reference Point

We consider a wide range of specifications of the reference point based on the lagged expectation of earnings. A model in which the reference point updates sufficiently slowly that it does not vary within our sample period could represent an ad-hoc model of daily income targeting in which the marginal utility of income declines substantially around the level of average daily earnings (Camerer et al., 1997). The first column of Table 4 considers this case by imposing a fixed reference point equal to the driver's average earnings across all shifts. Alternatively, we could characterize how expectations evolve throughout the day by estimating Equation (1) with total fare earnings for the shift as the outcome variable. The

Table 4—Maximum likelihood estimates: Alternative specifications

	(1) Fixed Ref. Point	(2) Previous Trip	(3) Loss Aversion: Hours	(4) Diminishing Sensitivity	(5) Stochastic Ref. Point
Disutil. of effort ψ	0.2722	0.2663	0.2980	0.0554	0.0070
	(0.1500)	(0.1401)	(0.1000)	(0.0267)	(0.0017)
Elasticity ν	0.8829	0.8877	0.1544	1.6678	2.2406
	(0.2576)	(0.2563)	(0.1545)	(0.2729)	(0.0715)
Loss aversion Λ	2.7228	2.6650	1.9180	2.7715	3.8762
	(0.3528)	(0.4043)	(0.1488)	(0.2832)	(0.1650)
Adaptation θ			0.7417	0.7896	0.6372
			(0.1226)	(0.0732)	(0.0137)
Std. dev. target ς					0.3839
					(0.0265)
Error term σ	0.6659	0.6616	0.2874	0.5404	0.4106
	(0.0856)	(0.1077)	(0.0526)	(0.0763)	(0.0154)
Mean log likelihood	-27,298	-27,312	-26,835	-26,903	-26,732
Test $\Lambda = 1$	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Test $\theta = 1$			< 0.001	< 0.001	< 0.001

Note: This table presents maximum likelihood estimates of Equation (6) with the adaptive reference point given by Equation (5). ς denotes the standard deviation of the target. The estimation sample consists of 34.5 million trips from over 37,000 drivers. The reference point in column (1) equals the driver's average earnings across all shifts. The reference point in column (2) equals expectations about earnings at the end of the previous trip. Column (3) corresponds to a model that incorporates loss aversion over hours. Column (4) corresponds to a model that incorporates diminishing sensitivity. Column (5) corresponds to a model with a stochastic reference point, in which the reference point consists of a distribution of earnings with variance ς. See Table 3 for further details.

second column defines the reference point as drivers' expectations at the end of the previous trip. We also consider reference points based on lagged expectations under different definitions of the lag (e.g., previous hour) and find relatively stable estimates, as Appendix Table 13 shows. In all cases, a likelihood ratio test rejects the null hypothesis of no loss aversion.

The reference points based on lagged expectations of earnings predict a pattern of income effects that do not match the observed recency pattern in the data. These specifications of the reference point create a stark contrast between the most recently accumulated earnings and any earlier earnings because the updated reference point incorporates fully the latter but not at all the former, as Appendix Figure 11 illustrates. The data instead show a gradual relationship between the timing of additional earnings and the probability of ending a shift.

Alternative Specifications of the Model

Section III.A makes several simplifying assumptions when introducing the model of reference dependence. We show that relaxing these simplifying assumptions does not change our conclusions about the importance of adaptive reference points.

First, the model focuses only on reference dependence in earnings. Kőszegi and Rabin (2006) posit that loss aversion operates in all dimensions of utility, and Crawford and Meng (2011) implement the model by including a reference point for hours in the cabdriver's objective function. Under this view, drivers experience losses from working longer than their "hours target," analogous to the losses from earning less than their "income target," with the same coefficient of loss aversion on both dimensions. An hours target seems particularly difficult to disentangle from neoclassical behavior because drivers may, for example, form commitments that coincide with their expectations about work hours. Nevertheless, we present results from a specification with a common coefficient of loss aversion for income and hours in Table 4 column (3). We also estimate a specification with a separate loss-aversion parameter for each dimension in Appendix Table 14. In both cases, incorporating loss aversion over hours substantially reduces the estimate of the elasticity parameter, but we obtain similar magnitudes for the remaining parameters.

Second, the gain-loss function does not exhibit the diminishing-sensitivity feature of prospect theory, and the income target I^r consists only of a point expectation. To incorporate diminishing sensitivity, we add curvature in the gain-loss function by calibrating a power function according to Tversky and Kahneman (1992). We find similar estimates for the main parameters of interest, though with a larger elasticity estimate and correspondingly smaller disutility of effort, as Table 4 column (4) shows. To estimate the model with a stochastic reference point that captures the distribution of potential earnings, we specify a normal distribution with a mean that updates as in Equation (5) and a variance given by the parameter ς. This specification yields a greater degree of loss aversion and correspondingly faster speed of adjustment of the reference point, as Table 4 column (5) shows. We also find a standard deviation of the income target of about \$38, along with a larger elasticity parameter and correspondingly smaller disutility of effort. Overall we find that relaxing the simplifying assumptions in the loss-aversion model does not change the conclusions about the importance of loss aversion and adaptive reference points.

IV. Discussion

The influential idea of income targeting informally captures the fact that accumulated daily earnings can decrease workers' willingness to continue working. However, evidence from NYC cabdrivers suggests that income targeting provides an incomplete account of daily labor-supply decisions.

A theory based on reference dependence helps to clarify the factors that govern daily labor-leisure tradeoffs. Beyond predicting a strong positive relationship between the probability of stopping work and accumulated hours worked much like the neoclassical model does, the reference-dependence model explains why we find only a weak positive relationship between stopping and accumulated income. Consistent with a reference point that takes time to adjust in response to new information, earlier experiences within the day gradually become incorporated into the reference point, thereby moderating the income effect, while recent experiences induce stronger behavioral responses.

Our framework suggests an important role for future work in understanding reference-point adaptation. Existing evidence covers a range of field settings that vary in the frequency of decisions, the size of the stakes, the familiarity of the choices, and the extent to which the decision maker pays attention to the problem. Early work by Simonsohn and Loewenstein (2006) documents history dependence in the willingness to pay for housing among movers between cities, which becomes offset for households that move again within their new city within a year. Post et al. (2008) illustrate the role of previous outcomes on risky choices in a game show, positing a limited form of reference-point updating based on changes relative to recent and initial conditions. DellaVigna et al. (2017) estimate the speed of adjustment of an income reference point on the scale of several months in the context of unemployment benefit cuts. Our setting enables us to shed light on reference-point formation by identifying the precise timing of reference-point effects. We contribute field evidence of high-frequency reference-point adaptation by focusing on a familiar, recurring decision problem in which the stakes consist of the workers' livelihood. Variation in the speed of adjustment across settings could arise due to differences in the frequency of decisions, the size of the stakes, the familiarity of the choices, and the extent to which the decision maker pays attention to the problem. Since the speed of adjustment governs the extent to which decision makers exhibit neoclassical behavior, a systematic characterization of reference-point adaptation would elucidate the sources and implications of reference-dependent behavior.

REFERENCES

Aalen, Odd O. 1989. "A Linear Regression Model for the Analysis of Life Times." *Statistics in medicine*, 8(8): 907–925.

Abeler, Johannes, Armin Falk, Lorenz Goette, and David Huffman. 2011. "Reference Points and Effort Provision." *American Economic Review*, 101(2): 470–492.

Agarwal, Sumit, Mi Diao, Jessica Pan, and Tien Foo Sing. 2015. "Labor Supply Decisions of Singaporean Cab Drivers." *Mimeo.*

Andersen, Steffen, Alec Brandon, Uri Gneezy, and John A. List. 2018. "Toward an Understanding of Reference-Dependent Labor Supply: Theory and Evidence from a Field Experiment." *Mimeo.*

Ashenfelter, Orley, Kirk Doran, and Bruce Schaller. 2010. "A Shred of Credible Evidence on the Long-run Elasticity of Labour Supply." *Economica*, 77(308): 637–650.

Banerji, Abhijit, and Neha Gupta. 2014. "Detection, identification, and estimation of loss aversion: Evidence from an auction experiment." *American Economic Journal: Microeconomics*, 6(1): 91–133.

Baucells, Manel, Martin Weber, and Frank Welfens. 2011. "Reference-Point Formation and Updating." *Management Science*, 57(3): 506–519.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2015. "Memory, Attention and Choice." *Mimeo.*

Bowman, David, Deborah Minehart, and Matthew Rabin. 1999. "Loss aversion in a consumption–savings model." *Journal of Economic Behavior & Organization*, 38(2): 155–178.

Buchholz, Nicholas, Matthew Shum, and Haiqing Xu. 2018. "Dynamic Labor Supply of Taxicab Drivers: a Semiparametric Optimal Stopping Model." *Mimeo.*

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): 407–441.

Cerulli-Harms, Annette, Lorenz Goette, and Charles Sprenger. 2019. "Randomizing Endowments: An Experimental Study of Rational Expectations and Reference-Dependent Preferences." *American Economic Journal: Microeconomics*, 11(1): 185–207.

Chang, Tom, and Tal Gross. 2014. "How Many Pears Would a Pear Packer Pack if a Pear Packer Could Pack Pears at Quasi-Exogenously Varying Piece Rates?" *Journal of Economic Behavior & Organization*, 99: 1–17.

Chou, Yuan K. 2002. "Testing Alternative Models of Labour Supply: Evidence from Taxi Drivers in Singapore." *Singapore Economic Review*, 47(01): 17–47.

Cleveland, William S., and Susan J. Devlin. 1988. "Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting." *Journal of the American Statistical Association*, 83(403): 596–610.

Crawford, Vincent P., and Juanjuan Meng. 2011. "New York City Cab Drivers' Labor Supply Revisited: Reference-Dependent Preferences with

Rational-Expectations Targets for Hours and Income." *American Economic Review*, 101(5): 1912–1932.

DellaVigna, Stefano. 2009. "Psychology and Economics: Evidence from the Field." *Journal of Economic Literature*, 47(2): 315–72.

DellaVigna, Stefano, Attila Lindner, Balázs Reizer, and Johannes F. Schmieder. 2017. "Reference-Dependent Job Search: Evidence from Hungary." *Quarterly Journal of Economics*, 132(4): 1969–2018.

Doran, Kirk. 2014. "Are long-term wage elasticities of labor supply more negative than short-term ones?" *Economics Letters*, 122(2): 208–210.

Dupas, Pascaline, Jonathan Robinson, and Santiago Saavedra. 2019. "The Daily Grind: Cash Needs, Labor Supply and Self-Control." *Mimeo.*

Ericson, Keith M. Marzilli, and Andreas Fuster. 2011. "Expectations as Endowments: Evidence on Reference-Dependent Preferences from Exchange and Valuation Experiments." *Quarterly Journal of Economics*, 126(4): 1879–1907.

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

Farber, Henry S. 2008. "Reference-Dependent Preferences and Labor Supply: The Case of New York City Taxi Drivers." *American Economic Review*, 98(3): 1069–1082.

Farber, Henry S. 2015. "Why You Can't Find a Taxi in the Rain and Other Labor Supply Lessons from Cab Drivers." *Quarterly Journal of Economics*, 130(4): 1975–2026.

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.

Frechette, Guillaume R., Alessandro Lizzeri, and Tobias Salz. 2016. "Frictions in a Competitive, Regulated Market Evidence from Taxis." *Mimeo.*

Gill, David, and Victoria Prowse. 2012. "A Structural Analysis of Disappointment Aversion in a Real Effort Competition." *American Economic Review*, 102(1): 469–503.

Gneezy, Uri, Lorenz Goette, Charles Sprenger, and Florian Zimmermann. 2017. "The Limits of Expectations-Based Reference Dependence." *Journal of the European Economic Association*, 15(4): 861–876.

Haggag, Kareem, and Giovanni Paci. 2014. "Default Tips." *American Economic Journal: Applied Economics*, 6(3): 1–19.

Haggag, Kareem, Brian McManus, and Giovanni Paci. 2017. "Learning by Driving: Productivity Improvements by New York City Taxi Drivers." *American Economic Journal: Applied Economics*, 9(1): 70–95.

Heffetz, Ori. 2018. "Are reference points merely lagged beliefs over probabilities?" National Bureau of Economic Research.

Heffetz, Ori, and John A. List. 2014. "Is the Endowment Effect an Expectations Effect?" *Journal of the European Economic Association*, 12(5): 1396–1422.

Imas, Alex. 2016. "The realization effect: Risk-taking after realized versus paper losses." *American Economic Review*, 106(8): 2086–2109.

Jonason, David, and Joel Wållgren. 2013. "Labor Supply of Stockholm Cab Drivers: Revisiting the Evidence." *Mimeo.*

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*, 47(2): 263–291.

Karle, Heiko, Georg Kirchsteiger, and Martin Peitz. 2015. "Loss Aversion and Consumption Choice: Theory and Experimental Evidence." *American Economic Journal: Microeconomics*, 7(2): 101–120.

Kőszegi, Botond, and Matthew Rabin. 2006. "A Model of Reference-Dependent Preferences." *Quarterly Journal of Economics*, 121(4): 1133–1165.

Kőszegi, Botond, and Matthew Rabin. 2009. "Reference-Dependent Consumption Plans." *American Economic Review*, 99(3): 909–936.

Martinussen, Torben, and Thomas H. Scheike. 2006. *Dynamic Regression Models for Survival Data.* Springer Science & Business Media.

Morgul, Ender Faruk, and Kaan Ozbay. 2014. "Revisiting Labor Supply of New York City Taxi Drivers: Empirical Evidence from Large-scale Taxi Data." *Transportation Research Board.*

Nguyen, Quang, and PinngSun Leung. 2015. "Revenue Targeting in Fisheries: The Case of Hawaii Longline Fishery." *Environment and Development Economics.*

Politis, Dimitris N., and Joseph P. Romano. 1994. "Large Sample Confidence Regions Based on Subsamples under Minimal Assumptions." *Annals of Statistics*, 2031–2050.

Post, Thierry, Martijn J. Van den Assem, Guido Baltussen, and Richard H. Thaler. 2008. "Deal or No Deal? Decision Making under Risk in a Large-Payoff Game Show." *American Economic Review*, 98(1): 38–71.

Schmidt, Marc-Antoine. 2017. "Identifying Behavioral Responses in Labor Supply: Idiosyncratic Shocks vs. Market-Level Shocks." *Mimeo.*

Simonsohn, Uri, and George Loewenstein. 2006. "Mistake #37: The Effect of Previously Encountered Prices on Current Housing Demand." *The Economic Journal*, 116(508): 175–199.

Song, Changcheng. 2016. "An Experiment on Reference Points and Expectations." *Mimeo.*

Sprenger, Charles. 2015. "An Endowment Effect for Risk: Experimental Tests of Stochastic Reference Points." *Journal of Political Economy*, 123(6): 1456–1499.

Stafford, Tess M. 2015. "What Do Fishermen Tell Us That Taxi Drivers Do Not? An Empirical Investigation of Labor Supply." *Journal of Labor Economics*, 33(3): 683–710.

Thakral, Neil, and Linh Tô. 2019. "Tipping and the Dynamics of Social Norms." *Mimeo.*

Tversky, Amos, and Daniel Kahneman. 1992. "Advances in Prospect Theory: Cumulative Representation of Uncertainty." *Journal of Risk and uncertainty*, 5(4): 297–323.

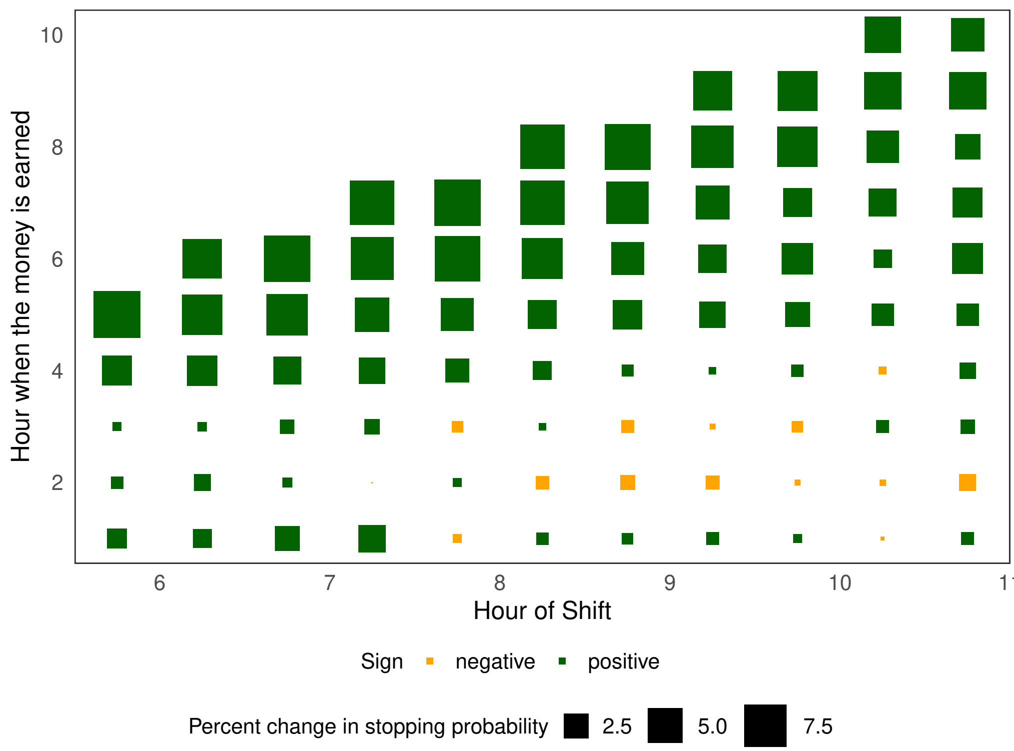

Figure 4. Stopping model estimates: Elasticity of stopping at different hours of the shift with respect to income—By timing of income

Note: The figure depicts the effect of an additional 10 percent in earnings accumulated at different times in the shift (vertical axis) on the probability of stopping at various times throughout the shift (horizontal axis). Each square has area proportional to the corresponding percent change in the probability of stopping. Estimates obtain from Equation (1) with the full set of controls (see Table 1 for details) and fixed effects for over 37,000 drivers. Each point represents estimates from trips that end within 5 minutes of the specified number of hours on the horizontal axis (ranging from 1.8 million trips to 0.4 million trips).