Daily Labor Supply and Adaptive Reference Points

Neil Thakral       Linh T. Tô*
Harvard University

May 2017

Abstract

We document evidence of high-frequency reference-point adjustment in the field. Analyzing a dataset of all New York City cab fares in 2013 using non-parametric methods, we find reductions in cabdriver labor supply in response to higher accumulated daily earnings and stronger effects for more recent earnings. The income effect is inconsistent with the neoclassical model and the non-fungibility of daily income rejects models invoking daily income targets. To explain the evidence, we incorporate adaptive reference points into models of loss aversion and salience. While loss aversion tends to overstate the main quantitative features of the data, both models capture the qualitative features.

*E-mail: nthakral@fas.harvard.edu and linhto@fas.harvard.edu. We are grateful to Isaiah Andrews, Pedro Bordalo, Stefano DellaVigna, Josh Feng, Edward Glaeser, Benjamin Golub, Jerry Green, Lawrence Katz, Marc Kaufmann, David Laibson, Matthew Rabin, Joshua Schwartzstein, Jesse Shapiro, Andrei Shleifer, and Elie Tamer for advice and suggestions.
1 Introduction

Reference dependence plays a contentious role in labor supply. Ever since Camerer et al. (1997) showed a negative relationship between average daily wages and the number of hours worked each day for taxi drivers in New York City, studies on the daily labor-supply decisions of workers who can flexibly choose their work hours have fallen into a clear dichotomy. The suggestion that workers may have a daily income target led to studies that either support the reference-dependence model if there is evidence of such a target, or uphold the neoclassical model with the rejection of an income target. This paper bridges these perspectives through the notion of adaptive reference points and quantifies the speed of reference-point adjustment.

We find two main empirical results by analyzing a dataset consisting of all New York City (NYC) cab fares in 2013 using non-parametric methods. First, we document excess sensitivity of labor-supply decisions to daily earnings. A driver who has already worked for 8.5 hours, for example, is on average 3 percent more likely to end a shift when cumulative daily earnings are 10 percent higher. This is contrary to the prediction of the neoclassical model that small windfalls should only trivially affect the marginal utility of lifetime wealth and therefore should leave labor-supply decisions unchanged. Second, we find stronger labor-supply reductions in response to earnings that accumulate more recently. One might naturally expect that the timing of income is economically irrelevant—an extra dollar earned at 12 PM is no different from an extra dollar earned at 11 AM from the perspective of a driver at 1 PM—or, if anything, recent earnings might be more informative about future earnings opportunities, which should make a driver less likely to end a shift in response to more recent earnings. However, we find that a driver who has already worked for 8.5 hours, for example, responds seven times more strongly to an additional dollar earned in the eighth hour of the shift compared to one earned in the fourth hour. The patterns persist for stopping decisions at different hours throughout the shift as well as for shifts that start at different hours of the day. We use an empirical Monte Carlo exercise to highlight the importance of the non-parametric methodology, as prior methods can yield spurious results due to functional-form assumptions.

We interpret these facts as evidence of a daily income effect and a violation of fungibility of money over time, even within a single day. We consider a number of alternative explanations including learning about future earnings, option value, and liquidity constraints, and show that these explanations cannot account for the patterns we observe. Furthermore, the income effect does not decrease with experience.
We also consider the possibility that daily earnings are correlated with unobserved determinants of labor supply such as effort or fatigue and show that the results hold when using instrumental-variables strategies that rely on variation in earnings due to distance or tips.

The findings are not only inconsistent with the neoclassical model but also with canonical behavioral explanations, necessitating an alternative formulation. Existing work, which invokes reference dependence and loss aversion to explain income-targeting behavior, does not account for the violation of fungibility for money earned at different times within a day. To the extent that a reference level influences decisions, our findings imply that the reference level must adjust within a day. A daily-level target for income does not permit stronger reactions to more recent earnings, and a reference level that adjusts instantaneously likewise does not make any distinction based on the timing of income, producing under some specifications behavior that resembles the neoclassical prediction. Our formulation consists of a slow-adjusting reference point, which incorporates experiences earlier in the day to a greater extent and results in stopping behavior that depends more strongly on recent earnings.

We explore several classes of models that can potentially explain the patterns of behavior that the data reveal. We develop structural models of daily labor supply and use the data on stopping decisions to estimate the parameters using maximum likelihood. As a benchmark, we consider models based on income targeting, expectations-based loss aversion (Kőszegi and Rabin, 2006), and salience (Bordalo et al., 2015), which can account for income effects but not the violation of fungibility within a day. We then propose two models that incorporate adaptive reference points, one based on loss aversion and one based on salience, and estimate the parameters of these extended models. We assess these explanations by comparing the estimated models’ predictions about the magnitude and timing pattern of the income effect, as well as by deriving and testing an additional prediction about how the magnitude of the income effect changes around the reference point. Both models capture the main qualitative features of the data, namely that labor supply reduces in response to earnings with stronger effects for more recent earnings. However, our model of loss aversion tends to overstate some of the quantitative features, as the maximum likelihood estimate of the coefficient of loss aversion implies a magnitude of the income effect and a change in behavior around the reference point that exceed what we observe in the data. We discuss how the results suggest an important role for stochastic reference points and provide support for models that predict a lower degree of loss aversion over money such as the notion of news utility in Kőszegi and Rabin (2009).
We highlight the implications of our findings for three different areas of work: daily labor supply, structural behavioral economics, and reference dependence.

The main empirical contribution of this paper is to provide evidence of daily income effects and a violation of fungibility by studying how workers adjust labor supply in response to small changes in accumulated daily earnings. By contrast, most studies in settings where workers can choose their own daily hours focus on the effects of transitory wage changes on labor supply. Camerer et al. (1997) present evidence of negative wage elasticities in two out of three samples of NYC cabdrivers, but the findings in this literature are mixed. Two recent papers revisit the setting of cabdrivers in NYC using a more comprehensive dataset: Morgul and Ozbay (2014) estimate elasticities for each of four months in 2013 using data from all trips taken in all taxi cabs in NYC and find a negative elasticity during only one of the months; Farber (2015) uses a sample of 13 percent of all cabdrivers between 2009 and 2013 and finds negative elasticities for only one-third of day-shift drivers and one-seventh of night-shift drivers. One potential concern with elasticity estimates in this setting is the implicit assumption that cabdrivers treat the average daily wage as parametric to their labor-supply decisions. Farber (2005) argues that the assumption of a parametric daily wage rate is unreasonable and instead proposes a model in which cabdrivers decide whether to stop working at the end of each trip. The stopping model serves as a starting point for our analysis, as we share the focus on daily income effects rather than elasticity estimates.

Despite the growing number of papers estimating high-frequency labor-supply elasticities, few authors investigate daily income effects. Using trip-level data for 21 cabdrivers, Farber (2005) concludes based on the stopping model that cumulative daily earnings do not significantly influence labor-supply decisions. Due to the limited sample, however, Goette et al. (2004) express concerns about the ability to identify income effects, which our paper addresses by using comprehensive administrative data. In addition, we document biases that arise due to functional-form assumptions in the stopping model (see Appendix C.2). We circumvent these biases using a non-

---


2Also see recent field experiments by Andersen et al. (2014) in India, which concludes that the labor supply of vendors in their Betel Nut Experiment remains unchanged in response to unexpected cash windfalls in the morning, and Dupas et al. (2016) in Kenya, which examines the effects of unexpected cash windfalls on a group of workers who have daily cash needs.
parametric methodology and, contrary to the result in Farber (2005), find evidence of substantial daily income effects. Moreover, we find that income effects depend on recency: a worker is disproportionately more likely to end a shift in response to a dollar received later in the shift than if the same dollar were received earlier in the shift.

As a second contribution of the paper, our model-comparison exercises provide a detailed comparison between different behavioral mechanisms that can potentially explain our findings. An emerging body of work in structural behavioral economics generally focuses on estimating the parameters of a single behavioral model to test the null hypothesis of the neoclassical model. Our structural models of daily labor supply serve several purposes: testing against the neoclassical model, testing whether reference points adjust, comparing the magnitudes of the income effects that the estimates predict, and testing an additional implication about moments of the data that the models were not designed to fit. Our results highlight that rejecting the neoclassical prediction does not imply a validation of a particular alternative. The few existing papers that analyze multiple behavioral theories generally reject the neoclassical model and broadly conclude support for the behavioral alternatives. Hastings and Shapiro (2013) find that consumers substitute to lower octane gasoline to an extent that cannot be explained by income effects when prices rise, consistent with their implementations of both loss aversion (Kőszegi and Rabin, 2006) and salience (Bordalo et al., 2013). Busse et al. (2014) analyze the effect of weather on car purchases and obtain results predicted by projection bias (Loewenstein et al., 2003) and also consistent with salience theory (Bordalo et al., 2013). Barseghyan et al. (2013) demonstrate the importance of probability weighting for explaining observed levels of risk aversion in insurance deductible choices, but cannot conclude whether loss aversion (Kőszegi and Rabin, 2006) or disappointment aversion (Gul, 1991) also plays a role. The present paper highlights the link between features of the data and assumptions of various behavioral models.

Prior work on daily labor supply tends to equate the question of whether workers behave according to the predictions of the neoclassical model with the sign of a single parameter, the elasticity of labor supply with respect to average daily wages, interpreting a positive elasticity as evidence in favor of the neoclassical model and a negative elasticity as evidence in favor of reference dependence and loss aversion. We depart from this paradigm for three reasons: first, within-day variation in the wage rate

---

3See, for example, Conlin et al. (2007), Crawford and Meng (2011), DellaVigna et al. (2012), Grubb (2012), and Laibson et al. (2015).
can bias elasticity estimates in either direction (see Appendix B); second, reference
dependence does not imply negative labor-supply elasticities if changes in wages
are anticipated (Kőszegi and Rabin, 2006); third, simply rejecting the neoclassical
prediction does not imply a validation of a particular alternative theory.

A third contribution of the paper is to provide field evidence on reference-point
formation and adjustment. A growing body of work shows that reference dependence
influences behavior in a variety of settings.4 However, much of the existing evidence on
how reference points adjust comes from decisions in game shows and lab experiments
(Post et al., 2008; Gill and Prowse, 2012; Song, 2016). More recently, DellaVigna et
al. (2017) find evidence of reference-point adjustment in the context of job search
over a horizon of several months. By contrast, the use of high-frequency labor-supply
data allows us to detect the speed of reference-point adjustment at the daily level.
Moreover, while the model in DellaVigna et al. (2017) involves a reference-dependent
utility function with a backward-looking reference point, a number of recent papers
provide evidence for forward-looking expectations-based reference points.5 We take
expectations-based reference points as a benchmark and express updated reference
points as a function of lagged expectations. Our formulation remains consistent with
the notion of reference points as recent expectations in Kőszegi and Rabin (2006)
but generates some backward-looking features, which can also account for empirical
results that find an influence of past prices on behavior in various settings (Odean,

The paper proceeds as follows. The next section provides background information
on the institutional context and describes the data. Section 3 analyzes the impact of
daily earnings on labor supply and discusses some possible explanations for the income
effect. Section 4 structurally estimates alternative classes of models and discusses the
implications of each. Section 5 concludes.

---

4See, for example, Barberis et al. (2001), Fehr and Goette (2007), Card and Dahl (2011), and
Allen et al. (2016).

5For example, see lab evidence from Abeler et al. (2011), Gill and Prowse (2012), Karle et al.
(2015), and Sprenger (2015) as well as field evidence from Post et al. (2008), Card and Dahl (2011),
2 Data

2.1 Background

Our study uses trip-level data provided by the New York City 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 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).6

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 had implemented a series of technology-based service improvements—including credit/debit card payment systems, passenger information monitors, text messaging between TLC and drivers, and automated trip sheet data collection—due to a March 2004 mandate by the Board of Commissioners of the TLC. Along with these service improvements, the automated trip sheet data include Global Positioning System (GPS) coordinates for pick-up and drop-off locations, which are available for over 98 percent of the data.

In each trip, the fare is determined by the meter. The standard city rate is calculated by combining a base rate of $2.50, any surcharges, and an additional amount that depends on the distance/time driven.7 Prior to September 4, 2012, the fare increases by $0.40 per unit (approximately 0.2 miles); afterwards the incremental charge is $0.50 per unit.8

Figure 1 depicts the average number of cabs that are on the road working during each minute of the day. The systematic drops in the number of cabs available in the early morning and early evening reflect the common institutional arrangement

---

6Haggag and Paci (2014) use the TPEP data in a study on default tip suggestions, which analyzes fares that were paid with a credit card and not subject to any surcharges. Farber (2015) also uses these data in a replication and extension of Camerer et al. (1997).

7The TLC reports three categories of surcharges: New York State Tax Surcharge of $0.50 for every ride on or after November 1, 2009; Night Surcharge of $0.50 between 8 PM and 6 AM; and Peak hour Weekday Surcharge of $1.00 between 4 PM and 8 PM. Source: http://www.nyc.gov/html/tlc/html/passenger/taxicab_rate.shtml.

8More precisely, the TLC defines a unit as follows: driving one-fifth of one mile at a speed of at least 6 miles per hour; or 60 seconds when the cab is not in motion or is traveling at less than 12 miles per hour. Source: http://www.nyc.gov/html/tlc/html/passenger/taxicab_rate.shtml.
whereby two drivers share the same cab (typically switching at 5 AM and 5 PM). 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.

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.

2.2 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. As in Haggag and Paci (2014), a shift is a sequence of consecutive trips that are not more than six hours apart from each other. In other words, a given trip is the last one in its shift if and only if it is followed by a period of at least six hours during which the driver does not pick up any more passengers. As in Farber (2005), we define a break as a long waiting time between fares: at least 30 minutes between a fare that ends in Manhattan and a fare that starts in Manhattan (over 85 percent of trips); 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 fare. After eliminating shifts with missing or inconsistent information (described in Appendix A), about 75 percent of the observations (over 5.8 million shifts by over 37,000 drivers) remain, comprising over $1.5 billion in transactions for cab fares.

Table 1 provides summary statistics at the trip level. Over 85 percent of all trips start and end in Manhattan, and the median ride is 10 minutes long. The median fare is about $9.5, with 90 percent of fares falling below $22. We observe tips for the 54 percent of fares that are paid using a credit card. A driver collects an average of $2.48 in tips per trip, but there is substantial variation in the rate of tipping.

---


10During 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. Source: http://www.nyc.gov/html/tlc/downloads/pdf/lease_cap_rules_passed.pdf.
as Figure 2 shows. Given any fare between the minimum fare of $2.5 and $60, the associated tip can take any value between $0.50 and $15, with higher concentrations of rounded tips or fixed fraction of the fare, anywhere between 10 percent and 35 percent. Around 65 percent of shifts contain a tip of at least $5, and 20 percent of shifts contain a tip of between $10 and $20. Haggag and Paci (2014) provide further evidence on variation in the rate of tipping.\footnote{For cabs that are equipped with credit-card machines from the largest vendor (accounting for 50 percent of cabs in NYC), there is a discontinuity in suggested tips when the fare reaches $15; Haggag and Paci (2014) exploit this discontinuity to show that default suggestions influence passenger tipping behavior.}

Figure 3 displays the fraction of shifts starting at each hour of the day as well as the distribution of work hours. A typical shift consists of 22 trips, with 75 percent of shifts exceeding 7.2 hours. About 64 percent of the time in an average shift is spent with a passenger in the cab, 26 percent of the time is spent searching for the next passenger, and 13 percent of the time is spent on break.

The market wage varies considerably throughout the day. For each minute that a driver spends searching for or riding with passengers, we define the driver’s per-minute wage as the ratio of the fare they earn for that trip to the number of minutes spent working (i.e., searching and riding). We define the market wage in each minute as the average of the per-minute wages of all drivers working during that minute. Figure 4 depicts the average wage during each minute of the day, and Appendix Figure 1 shows how the wage pattern differs between weekdays and weekends.\footnote{Appendix B documents that the pattern of wages throughout the day can be a source of bias for elasticity estimates.} The highest wages occur during the two hours with the lowest number of drivers, i.e., the transitions between AM and PM shifts each day (see Figure 1).\footnote{This corroborates the criticism due to Goette et al. (2004) that controlling for clock-hour effects removes much of the variation in earnings and quitting in the data from Farber (2005) because wages are highest precisely during the hours when many cabdrivers are required to end their shifts.} A cabdriver earns an average wage of about $31 per hour, which amounts to a gross income (excluding tips) of about $280 per shift, from which drivers may pay leasing fees and gasoline costs.\footnote{A medallion-cab lessor may agree to provide gasoline to drivers at no more than $21 per shift (or $126 per week). Source: \url{http://www.nyc.gov/html/tlc/downloads/pdf/lease_cap_rules_passed.pdf}.} Figure 5 investigates the predictability of hourly wages. 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), which represent a substantial component of the variation in hourly wages but do not reflect transitory daily variation, the figure shows a positive autocorrelation.
3 Tests of Income Effects

A neoclassical model of intertemporal utility maximization predicts that daily hours of work do not respond to small changes in accumulated daily earnings. As Appendix C.1 discusses in more detail, we model the decision of a driver at the end of each trip to stop working or to continue working. After completing \( t \) trips in \( h_{i,n,t} \) hours, driver \( i \) decides to end shift \( n \) when the disutility of effort for completing an additional trip outweighs the expected fare.\(^{15}\) A key prediction of the model is that there are no daily income effects: cumulative daily earnings \( y_{i,n,t} \) do not affect the decision to end a shift.

Letting \( d_{i,n,t} \) indicate the decision to stop working, we test the prediction that daily income effects are inconsequential by expressing the probability that driver \( i \) ends shift \( t \) at trip \( n \) non-parametrically as

\[
Pr(d_{i,n,t} = 1) = \sum_{j} \left[ (f_j(h_{i,n,t}) + \gamma_j y_{i,n,t} + X_{i,n,t} \beta_j + \mu_{i,j}) 1_{h_{i,n,t} \in H_j} \right] + \epsilon_{i,n,t}, \tag{1}
\]

where \( X \) consists of controls for location, time, and weather which can potentially be related to variation in earnings opportunities from continuing to work; \( f(h) \) is a function of work hours; \( \mu \) absorbs differences in drivers’ baseline stopping tendencies; and \( H_j \) partitions the minutes of the shift into intervals to allow a time-varying relationship between each of the covariates and the probability of stopping. The parameters \( \gamma_j \) capture the effect of an additional dollar on the probability of stopping during the \( j \)th time interval 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 3.4 discusses in more detail. The model predicts that \( \gamma_j = 0 \) for all \( j \), i.e., that the decision to end a shift is unrelated to cumulative daily earnings. Incorporating positively correlated hourly wages as in Figure 5 would yield the prediction that \( \gamma_j \leq 0 \), as greater recent earnings may indicate higher continuation values.

By interacting all covariates with a fine partition of the minutes in the shift, the regression does not impose functional-form assumptions that constrain the relationship between stopping and its potential determinants. For instance, this enables the

\(^{15}\)This assumption follows Farber (2005, 2008) and Crawford and Meng (2011), who suggest that not explicitly modeling option value is behaviorally reasonable. Moreover, for convex disutility of effort, this trip-by-trip stopping rule is consistent with maximizing the static objective function as long as the wage rate \( y'_n(h_n) \) does not increase too rapidly. The pattern in Appendix Figure 1 suggests that the stopping model may not apply to PM shifts on weekends, a point that we revisit when discussing empirical results in Section 3.1.
relationship between hours and the probability of stopping to be driver specific, whereas a parametric model with driver fixed effects would force 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. As another example, a standard fixed-effects model might suggest that 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 the non-parametric formulation allows the marginal effect of each variable on the probability of stopping to vary flexibly throughout the shift. Appendix C.2 conducts an empirical Monte Carlo exercise that validates this approach and demonstrates that stopping models estimated in prior work can yield spurious results.

3.1 Estimation of the Stopping Model

This section evaluates the prediction of the neoclassical model that cumulative daily earnings do not affect labor supply decisions. We partition the shift into 10-minute intervals and estimate the stopping model using Equation (1).\textsuperscript{16} Table 2 presents the results, with each row corresponding to a more comprehensive set of controls than the previous one. To interpret the magnitude of the income effect, the table reports the marginal effect of a 10 percent increase in cumulative earnings on the probability of ending a shift at 8.5 hours, which is approximately the median stopping time. The estimates in column (1) use 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). Columns (2) and (3) contain estimates from an alternative estimation strategy that uses distance between pick-up and drop-off locations to instrument for earnings, which we discuss further in Section 3.4.

All specifications consist of controls for minutes spent working, including indicators for the number of minutes in each hour with passengers. The specification in row 1 with no additional controls shows that higher cumulative daily earnings are associated with greater stopping probabilities, contrary to the prediction of the neoclassical model. Row 2 shows that using within-driver variation in earnings only strengthens the estimated effect. This contrasts with the results from Farber (2005), in which the positive effect of cumulative daily earnings on the probability of ending a shift becomes insignificant after accounting for interdriver differences in stopping probabilities.

\textsuperscript{16}All of the results are unchanged if we instead partition the shift into 30-minute intervals.
Section 3.3 discusses and compares the magnitudes in more detail.

In the remaining specifications, we include precisely measured controls to account for additional factors that can potentially affect labor supply and find that the income effect persists. To address the possibility that a driver is more likely to end a shift when a trip ends in a convenient location (e.g., near the driver’s home, or near a location where the cab can be transferred to another driver), the specification in row 3 includes indicators for the 195 Neighborhood Tabulation Areas (NTA) in NYC where a trip may end and an indicator for being in the zip code where the cab must be returned interacted with hour of the day.\textsuperscript{17} Figures 1 and 4 show a systematic pattern that many shifts end in the early evening, coinciding with a period of higher average wages, which reflects an institutional feature of the market that AM-shift drivers transfer shared cabs to PM-shift drivers and are therefore unable to serve passengers during that time. The estimates in row 4 include an interaction between clock hour and day of week as well as indicators for week of year and federal holidays. Row 5 uses high-frequency data from five stations in the NYC area to account for variation in stopping due to weather conditions. Farber (2015) points out that rainfall reduces the number of cabs on the street due to added disutility of driving in the rain. More generally, since adverse weather conditions can affect the labor-supply decisions of cabdrivers, we include the following indicators measured in the minute when a trip ends: precipitation, wind speed on the Beaufort scale, temperature above 80 degrees Fahrenheit, temperature below 30 degrees Fahrenheit. Under the full set of controls, a 10 percent increase in cumulative earnings corresponds to about a 3 percent increase in the probability of ending a shift (0.4 percentage-point increase relative to a baseline stopping probability of 13.6 percent) at 8.5 hours. Across all specifications, the data reveal a clear pattern of significant labor-supply reductions in response to cumulative daily earnings.

Figure 6 shows using the full specification (with driver fixed effects and controls for location, time, and weather) that the magnitude of the income effect from Table 2, evaluated at 8.5 hours of work, persists throughout the shift. The figure plots the income effect and probability of stopping every thirty minutes over a five hour period, roughly corresponding to the 10\textsuperscript{th} and 90\textsuperscript{th} percentile of the distribution of stopping times. As the average stopping probability varies from 4 percent to 27 percent, the magnitude of the income effect increases accordingly. Appendix Figure 7 plots the

\textsuperscript{17}We map the GPS coordinates of the drop-off location of each trip to the NTA, which is an aggregation of the census tracts. Source: http://www.nyc.gov/html/dcp/html/census/nyc_cff_faqs.shtml.
percent change in the probability of stopping estimated on four separate groups of shifts: day-weekday shifts, day-weekend shifts, night-weekday shifts, and night-weekend shifts.\textsuperscript{18} While the day shifts and the night-weekday shifts exhibit positive income effects consistent with our estimates from the full sample, the night-weekend shifts stands out with significant negative magnitudes. As Appendix C.1 points out, the trip-by-trip stopping model relies on the assumption that the option value of continuing to drive is sufficiently small (or that drivers ignore option value), and the pattern in Appendix Figure 1 suggests that this assumption may not be reasonable for night-weekend shifts, when wages rise substantially and predictably over time. This observation explains a potential discrepancy with the results in Farber (2015) that hours of work are roughly unaffected by income during night shifts.

3.2 The Role of Timing

This section provides a test of the fungibility of money within a shift by evaluating whether daily labor supply responds to the timing of earnings. The previous section tests for daily income effects but implicitly makes a standard economic assumption of fungibility, which in this case entails that the effect of income on the probability of stopping at any point during the shift depends only on cumulative earnings and not on how recently the dollars are received within a shift. We relax the assumption that the probability of stopping does not depend on the timing of earnings by augmenting Equation (1) to express the probability of stopping as

\[
\Pr(d_{i,n,t} = 1) = \sum_j \left[ \left( f_j(h_{i,n,t}) + \gamma_{j,k} \sum_k y_{i,n,t,k} + X_{i,n,t} \beta_j + \mu_{i,j} \right) \mathbb{1}_{\{h_{i,n,t} \in H_j\}} \right],
\]

where \(y_{i,n,t,k}\) denotes earnings accumulated in hour \(k\) of the shift. If money is fungible throughout the shift, then the impact of an additional dollar on the probability of ending a shift would be independent of when the dollar is earned (i.e., that \(\gamma_{j,k}\) is independent of \(k\)). A hypothesis based on the pattern of positively autocorrelated hourly wages in Figure 5 would be that \(\gamma_{j,k}\) is decreasing in \(k\) since greater recent earnings indicate higher continuation values.

Figure 7 plots the estimated effect of an additional $10 earned at various times in the shift on the probability of ending a shift at 8.5 hours using the full set of controls. The estimates provide evidence against fungibility: the effect that an additional dollar

\textsuperscript{18}Following Farber (2015), we define a day shift as one that starts between 4 AM and 10 AM and a night shift as one that starts between 2 PM and 8 PM. We classify night shifts on Friday and Saturday as well as day shifts on Saturday and Sunday as weekend shifts.
has on labor supply depends on how recently the dollar is received within a shift. The effect of an additional dollar on the probability of stopping is greater if the dollar is earned more recently. The magnitude is substantial, with an additional dollar accumulated one hour earlier increasing the probability of ending a shift by seven times more than an additional dollar accumulated five hours earlier. Based on the positive autocorrelation in earnings from Figure 5, higher recent earnings should, if anything, be associated with a higher value of continuing to work. However, holding total earnings fixed, the probability of ending a shift at the end of a given trip depends on the path of earnings throughout the shift in the opposite direction of this prediction: a driver is more likely to stop working after accumulating earnings at a higher rate toward the end of a shift. While additional earnings accumulated within the first three hours of the shift do not significantly impact the probability of stopping at 8.5 hours, drivers appear to respond strongly to income earned more recently in the shift by reducing labor supply.

Figure 8 shows that the violation of fungibility from Figure 7, evaluated at 8.5 hours of work, persists throughout the shift. The columns of the figure correspond to different times during the shift between hour 6 and hour 11. Within a given column (i.e., fixing a point in time during the shift), each row depicts the effect of an additional $10 accumulated in a particular hour on the probability of stopping: additional earnings accumulated early in the shift do not substantially affect the probability of stopping, but drivers are consistently more likely to end a shift in response to more recent earnings. The result that the timing pattern holds throughout the shift suggests an interpretation based on recency. The estimates in Figure 9 further confirm that time-of-day effects do not drive these patterns. The figure separates shifts into four groups based on start hour: morning (4 AM–10 AM), afternoon (10 AM–4 PM), evening (4 PM–10 PM), and night (10 PM–4 AM). Each column depicts the effect of an additional $10 earned at various times in the shift on the probability of ending a shift at 8.5 hours for a different group of shifts, demonstrating a timing pattern consistent with our estimates from the full sample.

3.3 Discussion of Magnitudes

To provide a better understanding of the size of the income effect, consider the estimates in Table 2. A 10 percent increase in cumulative earnings (an average of $26.67) corresponds to a 2.8 percent increase in the probability of ending a shift at 8.5 hours under the baseline specification with the full set of controls. If the additional
earnings arrive in the eighth hour of the shift, then the estimates in Figure 8 imply a 10.12 percent increase in the probability of stopping. For comparison, an additional 10 minutes of work (i.e., the median trip duration) increases the probability of ending a shift by 5.96 percent.

Most investigations of daily labor-supply decisions focus on the question of how workers adjust labor supply in response to wage changes. However, the few studies that examine daily income effects generally find that cumulative daily earnings do not affect labor supply.\(^\text{19}\) Farber (2005) finds an insignificant effect of earnings, though the point estimate implies that a 10 percent increase in cumulative earnings corresponds to an increase in the probability of ending a shift of 1.2 percent, which is less than half of the magnitude we find.\(^\text{20}\)

We find not only that cumulative daily earnings influence labor-supply decisions but also that the income effect depends on recency. The impact of a dollar earned in the eighth hour of a shift is seven times greater than that of a dollar earned four hours earlier, which implies a violation of fungibility of money earned within a shift. Our results complement existing studies on mental accounting, which tends to focus on how consumers treat money from different sources, by demonstrating violations of fungibility based on earnings accrued at different times within a single day.\(^\text{21}\)

### 3.4 Alternative Explanations

The evidence shows substantial changes in labor supply in response to small changes in wealth, and the magnitude of the reduction in labor supply depends on the timing...
of changes in wealth. This section addresses potential challenges to the modeling assumptions by considering the possibility that cumulative earnings are correlated with unobserved determinants of the stopping decision such as effort or fatigue, or that cumulative earnings convey information about future earnings opportunities. We also assess whether the relationship between cumulative earnings and stopping arises due to other factors such as option value, liquidity constraints, and inexperience.

3.4.1 Effort

We interpret the increase in the probability of ending a shift in response to higher cumulative earnings as evidence of a daily income effect exhibited by a reduction in labor supply. A potential concern with this conclusion is that labor supply consists of multiple dimensions, some of which are unobserved. If the reduction in hours coincides with an increase in the intensity of work, then the overall effect on labor supply would be unclear. We suggest three ways to address this: by using an instrumental-variable (IV) strategy, by constructing a proxy for effort, and by analyzing income effects and recency effects in the earlier hours of the shift.

First, columns (2) and (3) of Table 2 present IV estimates to address a possible correlation between cumulative daily earnings and unobserved determinants of the decision to end a shift such as effort or fatigue.\footnote{Although we do not report first-stage regressions, the F-statistics are sufficiently large for all of the IV specifications.} Since a cabdriver legally cannot refuse passengers because of destination, we instrument for cumulative earnings using the cumulative distance between pick-up and drop-off locations.\footnote{Camerer et al. (1997) argue based on a survey that driving passengers to a specific destination requires less effort than driving while searching for potential passengers.} The magnitude of the income effect does not significantly change, measuring distance either by GPS coordinates (column 2) or by odometer miles (column 3).\footnote{Since we control for the amount of time spent riding with passengers, variation in distance can reflect differences in driving speed. To alleviate concerns about a possible correlation between driving speed and unobserved determinants of the decision to end a shift, Appendix D.1 shows that the results hold using trips that stay within the dense streets of Manhattan, where variation in driving speed plausibly arises due to traffic conditions unrelated to the driver’s decisions to exert additional effort.} Additionally, Appendix D.1 provides an alternative IV strategy that exploits variation in income due to tips. To the extent that tips may be indicative of effort, a larger magnitude of the IV estimate might suggest an upward bias in the magnitude of the income effect. However, the IV analysis results in somewhat smaller estimates, consistent with our interpretation of the effect of cumulative earnings on the probability of stopping as representing an income effect. Moreover, estimating Equation (1) using either the distance or tip
instruments (see Appendix Figure 4) produces the same timing pattern of the income effect as in Figure 7, suggesting that the income effect does not reflect recent effort expended.

Second, although our dataset does not contain a direct measure of effort, we use as a proxy how quickly a driver finds the next passenger. Drivers spend 38 percent of their working hours searching for passengers during shifts in the bottom decile of earnings, compared with only 35 percent of working hours during shifts in the top decile. Part of the relationship reflects the mechanical fact that drivers likely earn more money during shifts that have a higher fraction of the time riding with passengers. Despite this mechanical effect, the data show only a weak relationship between earnings and the fraction of time spent searching (correlation of $-0.10$), suggesting a small role, if any, for adjustments along the effort dimension.\footnote{In addition, Figure 2 provides suggestive evidence of low returns to effort in terms of tips. Although the figure does not condition on any trip characteristics, the mass of points at a fare of exactly $52$ represents trips between Manhattan and JFK International airport and shows substantial variation in tips, ranging from $1$ to $20$.}

Third, Figures 6 and 8 show consistent and sizable income effects and recency effects throughout the shift, which poses a difficulty for fatigue-based explanations. An additional 10 percent in earnings corresponds to an increase in stopping probability of at least 2.5 percent with a significantly stronger response to recent earnings, even in the early hours of a shift. Furthermore, to the extent that drivers face an increasing marginal disutility of effort, we would expect much larger magnitudes of the income effect in the later hours of a shift (e.g., after working 10 hours compared to 8 hours) if effort poses a confound for estimating the effect of cumulative earnings. Using a non-parametric formulation mitigates the scope for the estimated income effect to reflect a response to fatigue, as Equation (1) allows for a driver-specific relationship between work hours and the probability of stopping and allows for a flexible relationship between the effect of work conditions on the probability of stopping.\footnote{Specifically, note that all controls and driver fixed effects interact with a fine partition of the minutes in the shift. This allows us to account for the effects of fatigue due to work hours and driving conditions as well as to accommodate neoclassical patterns of behavior such as quitting after a target number of hours without incorrectly attributing these to effects of income.}

### 3.4.2 Learning about future earnings

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

If higher cumulative earnings or higher recent earnings indicate lower expected
earnings from continuing conditional on all the covariates, then the estimated relationship between earnings and quitting would overstate the income effect. The pattern in Figure 5, however, suggests the opposite.

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. The evidence in Appendix D.2 suggests a limited role for this channel, as earnings do not appear predictive of market conditions on subsequent days.

3.4.3 Option value

The model in Appendix C.1 posits that at the end of each trip, drivers decide whether to end the shift or continue working for one more trip. In practice, however, a driver who believes that the wage will rise later in the shift might decide to continue working, and driver who explicitly solves the dynamic optimization problem might appear to have a low probability of ending a shift in response to low cumulative daily earnings. Neglecting option value would be problematic if the rate of increase in the wage exceeds that of the monetary equivalent of the disutility of effort (see Appendix C.1).27 Column (1) of Table 3 restricts the analysis to trips on Friday and Saturday after 5 PM, when the typical wage profile is nonincreasing (see Appendix Figure 1), and reports a positive relationship between earnings and stopping that does not significantly differ from the estimate using the full sample.

3.4.4 Liquidity constraints

Liquidity constraints often pose a challenge for identifying income effects. Johnson et al. (2006) and Parker et al. (2013), for example, find that household consumption exhibits excess sensitivity to small changes in wealth due to fiscal stimulus, but their results suggest an important role for liquidity constraints. Our work, by contrast, detects persistent income effects in labor-supply decisions at a high frequency, which limits the plausibility of an explanation based on liquidity constraints. Dupas et al. (2016) argue that bicycle-taxi drivers in Kenya set income targets as a commitment device to exert enough effort to meet daily needs, also pointing toward liquidity constraints. Such explanations in our setting would necessitate a consistent inability

---

27 As Figure 4 and Appendix Figure 1 show, wages typically do not rise sufficiently rapidly to justify concerns about this assumption, with the exception of night-weekend shifts which we highlight in Appendix Figure 7.
of NYC cabdrivers to smooth consumption across days.\textsuperscript{28} The result that drivers react differently to earnings accrued over different hours of the shift would be particularly difficult to rationalize based on liquidity constraints.

Although such effects are less plausible in our setting, we replicate 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.\textsuperscript{29} Although our data do not include information on ownership, we classify a driver as an owner-driver if (i) the driver operates exactly one cab, and (ii) no other driver shares that cab.\textsuperscript{30} The estimates in column (2) of Table 3 suggest that liquidity constraints do not confound the income effects we observe.

3.4.5 Experience

A hypothesis based on findings in related settings would be that the positive relationship between earnings and stopping reflects a failure to optimize by inexperienced drivers. Camerer et al. (1997) present evidence that more experienced drivers exhibit more positive wage elasticities of labor supply, which Farber (2015) corroborates. Recent work by Haggag et al. (2017) documents using the TPEP data from 2009 that productivity differences between new and experienced drivers vanish after 17 to 62 shifts (depending on the difficulty of the situations). Given that performance improves quickly with experience, drivers might also learn to supply labor more efficiently by ignoring daily earnings.

To consider the possibility of heterogeneity in income effects based on experience, column (3) of Table 3 restricts to the latest 10 percent of shifts in the sample for drivers with over 100 shifts. We find similar magnitudes of income effects as drivers gain more experience, with the full set of results that includes income effects at all deciles of experience in Appendix Table 4.

\textsuperscript{28}Camerer et al. (1997) argue that this seems unlikely because almost all lease-drivers pay their weekly fees in advance, and fleet drivers pay their daily fees at the end of the day or can pay late.


\textsuperscript{30}Our classification yields a subsample of owner-drivers, as we exclude those who lease to another driver. See Farber (2015) for additional institutional details on regulations concerning medallions in NYC.
3.4.6 Measurement error

At least two issues arise when measuring work hours in this setting. First, the data do not contain an explicit measure of break times. Second, 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. Appendix D.5 shows that accounting for the first issue does not change the magnitude of the income effect, and the second issue, if anything, biases the results against finding income effects.

4 Sources of Income Effects

The previous section provides evidence of excess sensitivity of labor-supply decisions to daily earnings, with a magnitude and timing pattern that a neoclassical model of labor supply cannot explain. This section estimates and compares alternative models which can potentially generate the income effects that the data reveal. In Section 4.1, we present a neoclassical model of labor supply as well as three behavioral models: income targeting, loss aversion, and salience. Section 4.2 extends the loss aversion and salience models to incorporate reference point adjustment, which we demonstrate is crucial for understanding the violation of fungibility shown in Section 3.2. The structural-estimation exercise in Section 4.3 yields two conclusions. First, the estimates reject the neoclassical model against several alternatives. Second, the models with adaptive reference points provide a better fit of the data than the corresponding models with static reference points.

Existing work that uses structural estimation to test a behavioral theory largely focuses on the objective of testing the null hypothesis of the neoclassical model, with a rejection of the neoclassical model interpreted as evidence in favor of a particular behavioral alternative. Section 4.4 departs from that paradigm by using the structural estimates to compare the models’ implications with the goal of adjudicating between the behavioral mechanisms. First, we compare the magnitude and timing pattern of the income effect predicted by each of the models. Second, we derive and test an additional prediction about moments of the data that the models were not designed to fit, namely how the magnitude of the income effect changes around the reference point.
4.1 Behavioral Models of Daily Labor Supply

The neoclassical model in Appendix C.1 posits that the marginal utility of lifetime income—and hence labor supply—does not vary in response to small, within-day changes in wealth. We assume the objective function for a driver with earnings \( I_t \) and hours of work \( H_t \) takes the form

\[
v(I_t, H_t) = v_I(I_t) + v_H(H_t) \\
= I_t - \frac{\psi}{1 + \nu} H_t^{1+\nu}, \tag{3}
\]

where \( \psi \) parameterizes the disutility of work and \( \nu \) is the elasticity parameter.

The stopping decision depends on the next trip’s expected fare \( \mathbb{E}_t[f_{t+1}] \) and duration \( \mathbb{E}_t[h_{t+1}] \). A driver with earnings \( I_t \) and hours of work \( H_t \) at the end of trip \( t \) decides to end a shift if the driver expects that completing an additional trip results in lower utility:

\[
\mathbb{E}_t[v(I_{t+1}, H_{t+1})] - v(I_t, H_t) + \varepsilon_t < 0, \tag{4}
\]

where \( I_{t+1} = I_t + \mathbb{E}_t[f_{t+1}], H_{t+1} = H_t + \mathbb{E}_t[h_{t+1}] \), and \( \varepsilon_t \) are independent and normally distributed with mean zero and variance \( \sigma^2 \). While the objective function depends explicitly on cumulative daily earnings \( I_t \), quasi-linearity implies that cumulative daily earnings do not affect the decision to end a shift.

To explain the results in Section 3, a model must allow for non-trivial within-day changes in the marginal utility of income. We analyze three different behavioral distortions that can create such income effects: income targeting, expectations-based loss aversion, and salience. Each of the models we consider makes an implicit assumption of narrow bracketing, that decision makers evaluate utility at the daily level.

4.1.1 Income targeting

This section formulates an ad-hoc model of daily income targeting in which the marginal utility of income declines substantially around the level of average daily earnings. As Camerer et al. (1997) suggest, a model in which drivers dislike falling short of their target more than they like exceeding it provides one possible explanation for the concavity in utility necessary for daily income effects.

The following objective function introduces a parameter \( \alpha \geq 0 \) to allow for a change in marginal utility at the target \( T \):

\[
v^{TT}(I_t, H_t) = v(I_t, H_t) + \alpha(I_t - T) \mathbf{1}_{\{I_t < T\}}. \tag{5}
\]
We interpret this as a model of reference-dependent preferences with a fixed reference point.\textsuperscript{31} The model implies a coefficient of loss aversion—the ratio between utility from losses and gains—of $1 + \alpha$. Farber (2015) tests the neoclassical model against this alternative hypothesis.

### 4.1.2 Expectations-based loss aversion

The theory of reference-dependent preferences due to Kőszegi and Rabin (2006) provides the leading explanation in the literature for the mixed evidence on behavior in daily labor-supply decisions.\textsuperscript{32} In their 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. Kőszegi and Rabin (2006) assume that rational expectations endogenously determine the reference point, consistent with laboratory experiments and field evidence in various settings.\textsuperscript{33}

Our primary analysis of loss aversion involves a simplified version of the model from Kőszegi and Rabin (2006), in which the objective function of the driver takes the following form:

$$v^{LA}(I_t, H_t) = (1 - \eta)v(I_t, H_t) + \eta \sum_{x \in \{I, H\}} n(x_t | x^r_t),$$

where $I^r$ and $H^r$ denote the reference levels for income and hours (i.e., the driver’s expected earnings and hours for the shift), and the gain-loss utility is given by

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

where $\eta$ determines the relative weight on gain-loss utility, and $\lambda \geq 1$ 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 $\eta = 0$). Despite adding two parameters to Equation (3), the model consists of only one additional degree of freedom since $\eta$ and $\lambda$ are not separately identifiable: behavior depends only on the

\textsuperscript{31}DellaVigna (2009) describes a model of this form as a simplified version of prospect theory (Kahneman and Tversky, 1979) that incorporates reference dependence and loss aversion without diminishing sensitivity and probability weighting.

\textsuperscript{32}The survey by DellaVigna (2009) discusses this, but also see more recent work by Crawford and Meng (2011).

\textsuperscript{33}For example, see lab evidence from Abeler et al. (2011), Gill and Prowse (2012), Karle et al. (2015), and Sprenger (2015) as well as field evidence from Post et al. (2008), Card and Dahl (2011), Ericson and Fuster (2011), and Pope and Schweitzer (2011).
ratio between utility from losses and gains, namely

\[
L = \frac{(1 - \eta) + \eta \lambda}{(1 - \eta) + \eta} = 1 + (\lambda - 1)\eta.
\]

This formulation makes two simplifying assumptions about the gain-loss component of utility. First, 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 and hours for that particular shift. Second, 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. While these assumptions follow the implementation from Crawford and Meng (2011), we relax both of them in Sections 4.3 and 4.4.

The Kőszegi and Rabin (2006) model features two important differences from the income-targeting model. First, drivers experience losses from working longer than their “hours target,” analogous to the losses from earning less than their “income target.” In other words, the separable gain-loss function indicates that drivers exhibit loss aversion not only over income but also over effort, with the same coefficient of loss aversion \( \lambda \) on both dimensions. Second, utility depends on expectations, as the reference points vary across drivers and across days. Section 4.2 discusses expectations and the specification of the reference point in more detail.

4.1.3 Salience

We adapt a model of salience based on Bordalo et al. (2015) to daily labor-supply decisions. The model combines two elements: (i) an evoked set determines the choice context and hence the salience of each attribute (income and hours), and (ii) decision makers place greater weight on the more salient attribute.\(^{34}\)

In this model, context influences decisions by distorting the relative weights that a driver places on income and leisure. A decision problem brings to mind an evoked set of options, each with an associated level of availability. The availability-weighted average of the options comprising the evoked set determines the normal levels of

\(^{34}\)Bordalo et al. (2012) develop a theory of choice under risk in which decision makers overweight states that are more salient. Bordalo et al. (2013) extend this concept to riskless choice among goods with multiple attributes (e.g., quality and price), where consumers place more weight on more salient attributes, but take the evoked set as exogenous. Our formulation of salience follows Bordalo et al. (2015) which models the evoked set explicitly.
income and hours. The extent to which an attribute varies within the evoked set relative to the normal level determines the salience of that attribute. Drivers place greater weight on the more salient attribute—income or hours—of their decision problem. In describing the components of the model more formally, we start with the salience distortions, taking the normal levels of income and hours as given, and then address how to determine the normal levels.

The objective function consists of a weighted sum

$$v^S(I_t, H_t) = \sum_{x \in \{I, H\}} \frac{w(\sigma(x_t, x^n_t), \delta)}{\sum_{y \in \{I, H\}} w(\sigma(y_t, y^n_t), \delta)} v_x(x_t),$$

where the relative weight $w(\sigma(x, x^n), \delta)$ on the utility for a given attribute increases in the salience $\sigma(x, x^n)$ of that attribute, $x^n$ denotes the normal level of the attribute, and $\delta \leq 1$ parameterizes the degree of distortion. We adopt the continuous salience weighting function from Bordalo et al. (2013):

$$w(\sigma(x, x^n), \delta) = \frac{[1 + \sigma(x, x^n)]^{1-\delta}}{2},$$

where the case $\delta = 1$ embeds the neoclassical model without context dependence, and the salience function $\sigma(\cdot, \cdot)$ is a symmetric and continuous function that satisfies ordering and diminishing sensitivity conditions. The ordering condition requires that moving an attribute further apart from the normal level increases its salience. Diminishing sensitivity expresses the idea that increasing the normal level renders a given difference between an attribute and the normal level less salient.\(^{35}\) Section 4.2 discusses the importance of these two properties for explaining the pattern of income effects. For a continuous and symmetric salience function satisfying these properties, Bordalo et al. (2012, 2013) suggest

$$\sigma(x, x^n) = \frac{|x - x^n|}{|x| + |x^n|},$$

which Hastings and Shapiro (2013) also use in empirical work.

To complete the description of the model, we discuss how to specify the evoked set and availability, which determine the normal levels of income and hours. We assume that the evoked set consists of the choices, stop or continue, and that the availability $a_t$ of stopping at the end of trip $t$ depends on how much more the driver must earn to

\(^{35}\)Note that the salience model defines diminishing sensitivity relative to zero, whereas the loss aversion model refers to diminishing sensitivity relative to the reference point.
reach the income target. We then define the normal level of an attribute as the availability-weighted average of the level of that attribute from stopping or continuing:

\[
I^n_t = a_t I_t + (1 - a_t) I_{t+1} \\
H^n_t = a_t H_t + (1 - a_t) H_{t+1}
\]

where \(a_t\) increases in \(I_t\), decreases in \(I^*_t\), and lies between 0 and 1. Intuitively, as earnings accumulate up to and beyond the income target, the decision to stop becomes more typical and thus comes to the top of the driver’s mind. For estimation, we assume that the availability of stopping corresponds to the predicted probability of ending a shift based on \(I_t\) and \(I^*_t\) from a logistic regression. We isolate the channel by which earnings influences stopping behavior through \(I_t\) and \(I^*_t\), though in principle other factors could be included as well, which we discuss in Section 4.4.2.

The idea that choice context influences how decision makers weight different attributes of a decision problem appears in a number of recent economic models (e.g., salience (Bordalo et al., 2013), focusing (Kőszegi and Szeidl, 2013), and relative thinking (Bushong et al., 2016)) which take the choice context as a degree of freedom called the evoked set, consideration set, or comparison set. These models conceptually distinguish between the choice set and the evoked set but equate them when deriving predictions. If the normal levels were an unweighted average of the elements in the evoked set as in (Bordalo et al., 2013), then assuming that the two sets coincide would fail to produce the pattern of income effects from Section 3 because doing so imposes fungibility: in such a model, the behavioral distortion only depends on cumulative earnings, independent of the timing of earnings. One way to proceed would be to make ad hoc assumptions about which additional options enter the evoked set. Instead, we assume the evoked set consists of the choices stop and continue, but we put structure on the components of the evoked set using a notion of availability based on Bordalo et al. (2015).

---

36 This assumption is based on Bordalo et al. (2015), who posit that availability is a map from past experiences and objective probabilities into a weight that reflects what comes to the decision maker’s mind.

37 We obtain similar results if we use different functional forms.

38 To explain observed gasoline-grade choice in 2006–2009, Hastings and Shapiro (2013) estimate a model of salience in which the evoked set consists of the current choices along with the grades of gasoline at the national mean prices from one week earlier. To improve fit, they propose an extended salience model in which the evoked set consists of the current choices along with the three grades of gasoline at prices $1.00, $1.10, and $1.20.
4.2 Adaptive Reference Points

In this section, we discuss how each of the behavioral models can potentially account for the evidence in Section 3. With the exception of the income-targeting model, which takes each driver’s target to be fixed across all shifts, the predictions depend crucially on the reference level. As a starting point, we take each driver’s reference level for each shift to be their rational expectations of income and hours for that particular shift. Under this view, the models of expectations-based loss aversion and salience yield daily income effects but treat money as fungible within the shift. By allowing for reference points that adjust within a shift, both models can generate the violations of fungibility necessary to explain the timing pattern of the income effect.

Consider the case of reference points based on rational expectations that represent drivers’ steady-state beliefs about earnings and hours. Under this view, reference points can be thought of as drivers’ targets for income and hours formed at the daily level, as in Crawford and Meng (2011): reference points vary across but not within shifts. We denote these ex ante expectations of shift-level earnings and hours by $I_{r0}$ and $H_{r0}$. The model based on Kőszegi and Rabin (2006) generates daily income effects through loss aversion, whereby the marginal utility of income decreases once earnings exceed the reference point. The same mechanism applies in the income-targeting model. The model based on Bordalo et al. (2015) generates daily income effects through the diminishing-sensitivity property of salience: the driver perceives the value of an additional fare less intensely at higher levels of daily earnings. Each of these models, despite generating income effects, treats money as fungible within the shift. Given the static reference point, the behavioral distortions depend on cumulative daily earnings in a shift, without scope for recent earnings to have a stronger influence on stopping decisions.

Suppose reference points adjust instantaneously to new information about how much the driver will earn by the end of the shift. In this case, if hourly earnings exhibit no substantial within-day autocorrelation, the driver updates expectations about earnings for the shift by immediately incorporating any difference between realized and expected earnings. Predictions of the Kőszegi and Rabin (2006) model approach those of the neoclassical model as reference-point adjustment becomes instantaneous. Intuitively, since rational expectations about daily earnings fully adjust, deviations from expectations no longer bring cumulative daily earnings closer to or further from the reference point. While the salience model continues to predict income effects due to diminishing sensitivity, no timing pattern emerges because instantaneous adjustment does not create a distinction between earnings at different
Stronger income effects in response to recent earnings requires a slow-adjusting reference point. Letting \( \Delta_t \) denote the difference between realized and expected earnings in trip \( t \), we model the reference point in this case as a convex combination of the lagged reference point and the expectation that obtains under full adjustment:

\[
I^r_t = \theta I^r_{t-1} + (1 - \theta) \left( I^r_0 + \sum_{\tau=1}^{t} \Delta_{\tau} \right),
\]

(8)

where \( 0 \leq \theta \leq 1 \), with \( \theta = 1 \) corresponding to a static reference point and \( \theta = 0 \) corresponding to a reference point that adjusts instantaneously. The recursive formulation produces a reference point based on lagged expectations:

\[
I^r_t = I^r_0 + \sum_{\tau=1}^{t} (1 - \theta^{t+1-\tau}) \Delta_{\tau}.
\]

The expression highlights that the reference point for income 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. This generates a violation of fungibility under both loss aversion and salience. Under loss aversion, the gain-loss component of utility depends on the difference between earnings and its reference level, and the reference point adjusts to a lesser extent in response to more recent earnings. The same applies to availability in the salience model, reflecting the intuition that recent earnings bring stopping closer to the top of the driver’s mind, which makes leisure relatively more salient due to the ordering property of salience. The qualitative predictions of both models under a slow-adjusting reference point corresponds to the following intuition about reacting to surprises: unexpected earnings constitutes a surprise, but surprises wear out over time so that quitting depends to a greater extent on recent earnings.

\[39\] With a positive autocorrelation of hourly earnings as in Figure 5, higher recent earnings should if anything affect the reference point more strongly and hence affect the stopping decision less strongly. This implies a higher probability of ending a shift in response to less recent earnings.

\[40\] When the availability of stopping increases, the normal levels \( I^n \) and \( H^n \) move away from the levels of these attributes under continuing \( (I_{t+1} \) and \( H_{t+1} \)). By the ordering property, this increases the salience of both income and hours, with the salience of hours increasing to a greater extent because of convexity in the disutility of work hours.
4.3 Structural Estimation

We use the data on stopping decisions to estimate the models via maximum likelihood under alternative specifications of the reference point.\footnote{Recall, however, that the income-targeting model depends on a fixed reference point, the driver’s mean earnings.} Each behavioral model produces a stopping rule analogous to Equation (4), resulting likelihood functions of the following form:

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

(9)

As a benchmark, we consider the case of static reference points based on rational expectations, proxying for expectations using the sample average of income and hours by driver and day of week (excluding the current shift) as in Crawford and Meng (2011). Although we estimate the parameters jointly for each of the models, the following intuition describes the main sources of identification. The disutility of effort $\psi$ is primarily identified by variation in work hours, and the elasticity parameter $\nu$ is primarily identified by variation in expected wages from continuing. The behavioral parameter in each model ($\alpha$ for income targeting, $L$ for expectations-based loss aversion, and $\delta$ for salience) is primarily identified by variation in cumulative earnings, which explains why we do not estimate a single model that nests and tries to distinguish multiple behavioral mechanisms. For the purpose of estimation, as in Section 3, we allow the parameters to be time-varying to avoid imposing any particular functional form on the relationship between work hours and quitting behavior. In what follows, we report estimates that reflect behavior after 8.5 hours of work using the same sample as in Table 2.

In Table 4, columns (2) to (4) report estimates of the three behavioral models corresponding to the objective functions from Equations (5) to (7). Each of these models nests the neoclassical model from Equation (3) with no income effects, reported in column (1). A likelihood ratio test in each case rejects the null hypothesis of the neoclassical model (i.e., $\alpha = 0$, $L = 1$, and $\delta = 1$). The results in column (3) relate to the main analysis in Crawford and Meng (2011), which also consists of a model of loss aversion with static reference points.\footnote{Our analysis uses fewer parameters than Crawford and Meng (2011) because they also estimate a set of control variables. We account for the control variables from Section 3 by residualizing the income variable and verified that this choice does not affect any of the results.} The income targeting model in column (2) implies a smaller change in the marginal utility of income at the target level. To the extent that the income targeting model mis-specifies the reference point, we would
expect to underestimate degree of loss aversion; however, the model does not consist of an hours target and hence may mis-attribute loss aversion over hours to income. The model of expectations-based loss aversion appears to provide the best statistical fit of the data based on the log-likelihood, though the models are not nested and thus cannot be directly compared using this criterion.

Relaxing the assumption of a static reference point, we estimate models of expectations-based loss aversion and salience with adaptive reference points. We allow the reference point to vary within a shift by introducing a new parameter, the adjustment term $\theta$, and defining the reference level $I^r$ according to Equation (8). This specification nests the static reference point above ($\theta = 1$) as well as a reference point that adjusts instantaneously ($\theta = 0$). Both extreme cases correspond to the fungibility of money within a shift, with $0 < \theta < 1$ indicating a violation of fungibility. The speed of adjustment is primarily identified by variation in the timing of income. Table 5 reports results consistent with a violation of fungibility, highlighting the importance of within-day adjustments. The estimate of $\theta$ differs significantly from 0 and 1 in both models, and a likelihood ratio test rejects the restriction to a static reference point. While $\theta$ appears close to 1 in both cases, its magnitude depends on the definition of a period. 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 $\theta$. The remaining parameters do not substantially differ from their counterparts in Table 4.

The structural estimates for both models provide empirical support for the notion of adaptive reference points. One interpretation of the evidence suggests a departure from rational expectations as the reference point. Rational expectations derive from a steady-state distribution of potential earnings, and a given shift constitutes a particular realization from that distribution. This view of rational expectations predicts a static reference point ($\theta = 1$), which the evidence rejects. Under an alternative interpretation, rational expectations incorporate new information within a shift, which leads the reference point to adjust. While this does not require that expectations or beliefs adjust slowly to new information, the evidence does suggest that preferences do not instantaneously change, consistent with the idea from Kőszegi and Rabin (2006) that preferences depend on lagged expectations.

\footnote{The results in Appendix E also reject a reference point that does not adjust (i.e., $\theta = 1$) if we allow for stochastic reference points that capture the distribution of potential earnings and hours.}
4.4 Comparison of Models

This section provides a more detailed comparison between the models of expectations-based loss aversion and salience by exploring their implications for explaining the patterns of stopping behavior in the data.

4.4.1 Magnitude of income effect

This section assesses whether drivers who behave according to the estimates in Table 5 exhibit a relationship between earnings and quitting that conforms to the results from Section 3. Figure 10 shows the timing pattern of the income effect under the estimated models of expectations-based loss aversion and salience with adaptive reference points. For each model, the figure plots the predicted effect of an additional $10 earned at various times in the shift on the probability of ending a shift at 8.5 hours, with the estimates from Figure 7 providing a benchmark for comparison. As Section 4.2 highlights, adaptive reference points in each case lead to a violation of fungibility, with stronger labor-supply reductions in response to more recent earnings.

Although both models predict income effects that are qualitatively consistent with the pattern in Figure 7, the models make different predictions about the magnitude. The salience model from Equation (7) produces income effects that largely coincide with the reduced-form estimates. However, the loss-aversion model from Equation (6) leads to magnitudes that consistently exceed the income effects from the data. Re-estimating the model after relaxing the simplifying assumptions from Section 4.1.2 by allowing for stochastic reference points or curvature in the gain-loss function does not change this result (see Appendix E).

As a possible explanation for the large estimated magnitudes, note that the Kőszegi and Rabin (2006) model assumes a constant coefficient of loss aversion across different dimensions of utility. To test whether this affects the results, we relax the assumption implicit in Equation (6) to allow for a different coefficient of loss aversion on each dimension ($L_I$ for income and $L_H$ for hours). The assumption of a “universal gain-loss function,” designed to avoid introducing additional degrees of freedom, does not appear consistent with the estimates in Table 6. Although the estimates reveal a significant degree of loss aversion on both dimensions, a likelihood ratio test rejects the restriction that $L_I = L_H$. The degree of loss aversion over income is significantly smaller, consistent with the claim from Farber (2015) that the patterns in the data suggest a larger role for reference dependence in daily hours compared to income. The coefficients of loss aversion in Tables 4 and 5 attempt to simultaneously fit behavior
over both of these dimensions, thereby overstating the degree of loss aversion over income.

This observation highlights the importance of analyzing the reduced-form implications of estimated models. A likelihood ratio test would simply lead to the conclusion that the behavioral models provide a better fit for the data than the neoclassical model. This says little about the ability of the models to explain the patterns of interest, especially those that are not directly matched. The exercise of checking what the parameter estimates imply about the underlying behavior that the model intends to capture leads to richer conclusions beyond statistical measures of fit.

While much of the motivation for studying daily income targeting stems from the observation that under some circumstances workers appear to work less when they can earn more, various forces naturally push in the opposite direction, each of which poses challenges for uncovering the role of reference dependence in earnings. Higher expected wages lead drivers to work more both in the neoclassical model and in the reference-dependent model (Kőszegi and Rabin, 2006). Moreover, as Section 4.2 highlights, reference-point adjustment modulates the effects of loss aversion. Finally, if money constitutes news about future utility rather than contemporaneous consumption utility, then loss aversion over income can play a less pronounced role (Kőszegi and Rabin, 2009). We address the first by analyzing behavior through the stopping model and the second by explicitly evaluating the role of timing. Under the view that our estimate of loss aversion over income is attenuated by news utility, the results provide an even stronger demonstration of reference dependence, though a better understanding of the relative importance of loss aversion over monetary outcomes remains a topic for future research.

4.4.2 Change in behavior at the reference point

This section considers an additional implication of the behavioral models from Section 4.1. Loss aversion and salience posit different mechanisms by which the income effect arises. Loss aversion invokes a sharp change in behavior near the reference levels. The model in Section 4.1.2 predicts, holding fixed income and hours, that the effect of income on the probability of stopping changes significantly when income is below the target compared to when income is above the target. The model of salience in Section 4.1.3, on the other hand, predicts a smooth change as income varies. The income effect arises through diminishing sensitivity, leading the driver to undervalue additional fares at higher levels of income, without a substantial change near the income target.
To directly test these predictions about how reference points influence stopping behavior, we re-estimate the marginal effect of earnings on the probability of stopping at 8.5 hours by extending Equation (1) to allow the income effect to vary based on distance from the income and hours target. We consider the set of trips that fall within $60 of the income target and within 2 hours of the hours target and categorize trips into 144 groups: twelve 20-minute intervals based on distance from the hours target interacted with twelve $10 intervals based on distance from the income target. Figure 11 depicts the effect of an additional $10 for each of these groups, with rows representing distance from the income target and columns representing distance from the hours target. Within each column, holding fixed the distance from the hours target but varying the distance from the income target, the income effect appears relatively stable. The magnitude of the income effect increases gradually with distance from the hours target but does not vary with distance from the income target.

Analyzing stopping behavior near the target helps to evaluate the simplifying assumptions in the model of loss aversion from Section 4.1.2. Once the level of income exceeds its target, the effect of an additional dollar on the probability of stopping under the assumption of a piecewise-linear gain-loss function becomes zero, which does not hold in the data. Models based on prospect theory (Kahneman and Tversky, 1979; Kőszegi and Rabin, 2006) typically posit a value function that exhibits diminishing sensitivity (convexity in losses and concavity in gains), but convexity in losses would predict a negative marginal effect of earnings on the probability of ending a shift below the income target. Relaxing the assumption that \( I_r \) and \( H_r \) denote the driver’s point expectations for a given shift, a stochastic reference point (Kőszegi and Rabin, 2006) which represents the full distribution of a driver’s potential earnings for that particular shift yields a smoother relationship between the income effect and distance from the target. For a driver who experiences gain-loss utility relative to the full distribution of potential earnings, additional income constitutes a partial mitigation of loss whether accumulated daily earnings fall above or below the target level. Finding positive income effects that do not change with distance from the income target therefore suggests an important role for stochastic reference points.\(^{44}\)

The relationship in Figure 11 also clarifies the role of the assumptions in the salience model from Section 4.1.3. In the salience model, the availability of stopping increases continuously as income accumulates up to and beyond the income target.

\(^{44}\)The estimates in Appendix E of the loss-aversion model with stochastic reference points show a substantial improvement in the log-likelihood compared to the model with point expectations as the reference level.
For simplicity, our model of salience assumes that availability depends on income but not hours. Since the normal levels of both income and hours depend on availability, the marginal effect of earnings on the probability of stopping in the salience model varies with distance from the income target but not distance from the hours target. The pattern in the data that the magnitude of the income effect increases as the driver approaches or passes the hours target points toward hours as a key determinant of availability.

5 Discussion

This paper documents violations of fungibility for money earned at different times. Contrary to the neoclassical prediction, drivers treat an additional dollar on one day as being different from a dollar on another day, resulting in a daily income effect whereby labor supply decreases in response to accumulated daily earnings. Moreover, drivers are more likely to stop working in response to earnings accumulated more recently within the same day. These facts taken together are inconsistent with a notion of income targeting in which drivers reduce labor supply after earning a particular amount each day. However, models of reference-dependent preferences can explain both facts with a reference level that adjusts within the day. As reference levels take time to adjust in response to recent changes in expectations, earlier experiences within the day become incorporated into the reference point, thereby moderating the income effect, while recent experiences induce stronger behavioral responses. Our findings provide field evidence for reference dependence and reference-point adjustment that persist in the face of experience and high stakes in a familiar, recurring setting.

The mechanisms explored in our setting may also be relevant for labor supply decisions more generally, especially with the rise of alternative work arrangements (Katz and Krueger, 2016). Chen and Sheldon (2015) use data from the ride-sharing company Uber to study income-targeting behavior as in Camerer et al. (1997) and Farber (2005) and finds mixed evidence. While driver-partners are more likely to quit in response to higher cumulative earnings under some specifications, surge pricing tends to increase the supply of rides. A model of adaptive reference points could explain why cumulative earnings from fares may create a daily income effect while surge pricing generates the opposite behavior, as drivers’ reference points for income quickly adjust upward to reflect higher anticipated earnings from future trips due to surge pricing. Further research can investigate the role of reference dependence and adaptive reference points in flexible work relationships as self-employment and
contract work are becoming increasingly prevalent.

More broadly, the mechanisms behind the violations of fungibility could have implications for the design of government tax and transfer systems. For example, marginal propensities to consume may respond to the timing of income through institutional features such as tax withholding (Shapiro and Slemrod, 1995; Feldman, 2010). Violations of fungibility via the timing of income could also help create and understand effective policy instruments for encouraging retirement savings or providing fiscal stimulus (Shefrin and Thaler, 1988; Souleles, 1999). Future work can explore the influence of adaptive reference points in these as well as other field settings.

References


A  Data

As in Haggag and Paci (2014), we first process the data by dropping data errors including those resulting from electronic tests.

1. If the drop-off time is before the pick-up time in a trip, then we swap the drop-off time and pick-up time: 0.01 percent of the trips.

2. If the same driver or the same car’s drop-off time is after the pick-up time of a subsequent trip, then we set the drop-off time to be equal to the pick-up time of the subsequent trip: 0.06 percent of the trips.

We flag trips that have any of the inconsistencies outlined below:

1. Trips that have distance of zero: 0.65 percent of the trips.

2. Trips that have ride duration of zero: 0.3 percent of the trips.

3. Trips with payment type recorded as “No Charge” or “Dispute”: 0.42 percent of trips.

4. When fare is too high or too low compared to distance and time and locations: 0.16 percent of the trips.

5. When fare is too low compared to distance and time and locations: 0.25 percent of the trips.

6. Trip time as indicated by the pick-up and drop-off timestamps and the recorded ride duration do not match: 0.20 percent of the trips.

7. Outliers with trip durations longer than 3 hours or trip distance longer than 100 miles: 0.65 percent of the trips.

8. Trip between Manhattan and an airport in under 5 minutes: 0.07 percent of the trips.

9. Trip between Manhattan and JFK International airport in under 10 miles: 0.06 percent of the trips.

10. When a ride lasts fewer than ten seconds, or fewer than one minute and costs over $10: 0.68 percent of the trips.
11. When a ride lasts fewer than ten seconds, or fewer than one minute and costs over $10: 0.49 percent of the trips.

12. When average speed during a trip exceeds 80 miles per hour: 0.39 percent of the trips.

13. Trips belonging to truncated shifts (those that start before the first day or end after the last day of the year): 0.09 percent of the trips.

We then remove shifts with trips that have been flagged with errors or shifts that are outliers:

1. Trips in the same shift but with more than one car: 0.40 percent of the shifts.

2. Shifts that are longer than 18 hours: 1.12 percent of the shifts.

3. Shifts that are shorter than two hours: 2.31 percent of the shifts.

4. Shifts by drivers with under 100 rides on record (may be electronic tests sent by TLC or the vendors): 0.12 percent of the shifts.

5. Shifts with fewer than three trips: 1.5 percent of the shifts.

As many of the analyses require that trips are in successive order, we remove the whole shift when one trip is questionable, and the process reduces our sample by approximately 24 percent. After cleaning out shifts, we remove drivers with under ten shifts, or an additional 0.1 percent of the observations. We are left with a sample of 127 million observations from over 37,000 drivers in over 5.8 million shifts. Finally, in many of the analyses, we restrict our sample to shifts that stay within the five boroughs in NYC, consisting of 94 percent of the remaining shifts.

B Elasticity Estimates

In this section, we discuss how the wage profile can potentially lead to mechanical biases of wage-elasticity estimates. To illustrate how elasticities can be biased in the positive direction, consider a hypothetical driver who supplies labor inelastically, with some noise around the optimal stopping time. If the wage increases throughout each shift, then a regression of log hours on log wages will have a positive coefficient on log wages since longer shifts mechanically have higher average earnings.
We provide suggestive evidence that this bias might be present in the setting of cabdrivers in NYC by estimating elasticities on subgroups of shifts with different wage patterns. To estimate wage elasticities, we follow the approach by Camerer et al. (1997) and Farber (2015), regressing the logarithm of the total working hours in a shift on the logarithm of the average earnings per hour in that shift, with time controls (indicators for day of week, week of year, and federal holidays) and driver fixed effects. We also use average market wage of a non-overlapping sample of drivers to instrument for a driver’s wage.

Appendix Figure 1 displays the pattern of average wages on weekdays and on weekends. Though the patterns are similar across AM shifts, they diverge significantly between 10 PM and 1:30 AM when the average wage is rising for weekend shifts but falling for weekday shifts. As around half of the cabdrivers who work during the PM shift stop during this period, this distinction may have a nontrivial impact on the elasticity estimates. For each type of shift (day or night), we estimate the wage elasticity of weekday shifts and weekend shifts separately, restricting the sample to drivers who appear in both groups to avoid compositional differences in responsiveness to wage changes. The estimates in Appendix Table 1 confirm that while wage elasticities across weekdays and weekends are similar for AM shifts, they are substantially higher during weekends for PM shifts, consistent with the pattern of increasing average wages on weekend nights. Instrumental-variable estimates would imply striking differences in behavior between weekdays and weekends for PM shifts, with an elasticity of 0.3067 on weekends and 1.3085 on weekdays, and no such difference in behavior for AM shifts. As the direction of these estimates coincides with predictions based on daily wage patterns, our results suggest that within-day variation in wages can lead to biases in elasticity estimates.

C Model and Estimation

C.1 A Model of Daily Labor Supply

We begin by presenting a neoclassical model of intertemporal utility maximization with time-separable utility. We then formulate testable predictions about daily labor-supply decisions and proceed to evaluate these predictions using data on the labor supply of NYC cabdrivers.
An individual maximizes lifetime utility given by

\[ U = \sum_{t=0}^{N} \rho^n u(c_n, h_n), \]

where \( \rho \) is the discount factor, \( c_n \) is consumption in period \( n \), \( h_n \) is hours worked in period \( n \), and \( u(\cdot) \) is a per-period utility function which is increasing in consumption, decreasing in hours worked, and concave in both arguments. The lifetime budget constraint is given by

\[ \sum_{t=0}^{N} (1 + r)^{-1} (y_n(h_n) - p_n c_n) = 0, \]

where \( p_n \) denotes the price of consumption, \( r \) denotes the interest rate, and daily earnings \( y_n(\cdot) \) is an increasing function of labor supply. The first-order conditions for this intertemporal maximization problem equate the marginal utility of lifetime income with the marginal utility of consumption and the marginal disutility of effort per unit of wage.

The problem of maximizing lifetime utility is equivalent to that of maximizing a static one-period objective function

\[ v(h_n) = \lambda y_n(h_n) - g(h_n, \lambda p_n), \quad (10) \]

where the monetary equivalent of the disutility of effort \( g(\cdot) \) is convex and \( \lambda \) is the lifetime marginal utility of income along the optimal path.\(^{45}\) Taking one period to be a shift, we model the decision of a driver to continue working or to stop working at the end of each trip. To evaluate whether to continue working, drivers must form expectations about the additional income earned from and the additional time spent on a prospective trip.

A driver decides to end a given shift when the disutility of effort for completing an additional trip outweighs the expected fare.\(^{46}\) After completing \( t \) trips in \( h_{i,n,t} \) hours, driver \( i \) decides to end shift \( n \) if \( v(h_{i,n,t}) \) exceeds the value \( v(h_{i,n,t+1}) + \varepsilon_{i,n,t} \) of continuing to work for one more trip, where the error terms \( \varepsilon_{i,n,t} \) are independently drawn from a distribution \( F \). We define \( d_{i,n,t}^* = v(h_{i,n,t+1}) + \varepsilon_{i,n,t} - v(h_{i,n,t}) \) as the

\(^{45}\)See the online appendix of Fehr and Goette (2007) for a derivation following Browning et al. (1985).
\(^{46}\)This assumption follows Farber (2005, 2008) and Crawford and Meng (2011), who suggest that not explicitly modeling option value is behaviorally reasonable. Moreover, for convex disutility of effort, this trip-by-trip stopping rule is consistent with maximizing the static objective function as long as the wage rate \( y'_n(h_n) \) does not increase too rapidly.
latent value of continuing for another trip and let \( d_{i,n,t} = 1 \{ d_{i,n,t} < 0 \} \) indicate the decision to stop working. A key prediction of the model is that there are no daily income effects: cumulative daily earnings \( y_{i,n,t} := y_n(h_{i,n,t}) \) do not affect the decision to end a shift. Taking a reduced-form approximation for the value of continuing, we test the prediction that daily income effects are inconsequential by expressing the probability that driver \( i \) ends shift \( t \) at trip \( n \) non-parametrically as

\[
\Pr(d_{i,n,t} = 1) = \sum_j \left[ (f_j(h_{i,n,t}) + \gamma_j y_{i,n,t} + X_{i,n,t}\beta_j + \mu_{i,j}) \mathbf{1}_{\{h_{i,n,t} \in H_j\}} \right] + \epsilon_{i,n,t}, \tag{1}
\]

where \( X \) consists of controls for location, time, and weather which can potentially be related to variation in earnings opportunities from continuing to work; \( f(h) \) is a function of work hours; \( \mu \) absorbs differences in drivers’ baseline stopping tendencies; and \( H_j \) partitions the minutes of the shift into intervals to allow a time-varying relationship between each of the covariates and the probability of stopping. The model predicts that \( \gamma_j = 0 \) for all \( j \), i.e., that the decision to end a shift is unrelated to cumulative daily earnings.

### C.2 Simulation Exercise

To evaluate various approaches for estimating stopping behavior, we conduct a set of empirical Monte Carlo studies (Stigler, 1977; Huber et al., 2013). The data for our simulations consists of a sample of over 3 million trips from 1,000 drivers. The first set of simulations considers stopping decisions that do not depend on earnings. The second set of simulations considers stopping decisions that depend on cumulative daily earnings but not on the timing of earnings. We find that the non-parametric approach in the present paper produces the expected result across all of the simulations, whereas alternative approaches from the literature may yield a significant positive or negative effect of earnings (and the timing of earnings) on stopping.

We follow the notation in Appendix C.1, where \( d_{i,n,t} \) denotes the decision to stop working, \( y_{i,n,t} \) denotes cumulative earnings, and \( h_{i,n,t} \) denotes the number of hours driver \( i \) has worked at the end of \( t \) trips in shift \( n \). Letting \( \Phi \) denote the standard normal cumulative distribution function, we consider regression equations of the
following forms:

\[
\text{Pr}(d_{i,n,t} = 1) = \sum_j \left[ (\alpha_j h_{i,n,t} + \gamma_j y_{i,n,t} + \mu_{i,j}) \mathbf{1}_{\{h_{i,n,t} \in H_j\}} \right] + \epsilon_{i,n,t} \quad (\text{TT})
\]

\[
\text{Pr}(d_{i,n,t} = 1) = \Phi(\alpha h_{i,n,t} + \gamma y_{i,n,t} + \mu_i) \quad (F-1)
\]

\[
\text{Pr}(d_{i,n,t} = 1) = \Phi\left( \sum_j \alpha_j \mathbf{1}_{\{h_{i,n,t} \in H_j\}} + \gamma y_{i,n,t} + \mu_i \right) \quad (F-2)
\]

\[
\text{Pr}(d_{i,n,t} = 1) = \Phi\left( \sum_j \alpha_j \mathbf{1}_{\{h_{i,n,t} \in H_j\}} + \sum_j \gamma_j \mathbf{1}_{\{y_{i,n,t} \in \hat{Y}_j\}} + \mu_i \right) \quad (F-3)
\]

\[
\text{Pr}(d_{i,n,t} = 1) = \Phi\left( \sum_j \alpha_j \mathbf{1}_{\{h_{i,n,t} \in H_j\}} + \sum_{j,\ell} \delta_{j,\ell} \mathbf{1}_{\{h_{i,n,t} \in H_j\}} \mathbf{1}_{\{y_{i,n,t} \in \hat{Y}_\ell\}} + \mu_i \right) \quad (F-4)
\]

Equation (TT) corresponds to Equation (1) in the main text, excluding the control variables, with \( H_j \) partitioning the shift into 10-minute intervals. Equation (F-1) resembles the probit model from Farber (2005) and Crawford and Meng (2011), in which income and hours are constrained to enter linearly. Equation (F-2) relaxes the constraint by allowing for a non-parametric relationship between hours and the probability of stopping. Equation (F-3) corresponds to the alternative specification in Farber (2005) when we take \( \hat{H} \) and \( \hat{Y} \) to partition the shift at \{180, 360, 420, 480, 540, 600, 660, 720\} minutes and \{25, 50, 75, 100, 125, 150, 175, 200, 225\} dollars, respectively. The main specification in Farber (2015) corresponds to Equation (F-3) and the more flexible specification in Farber (2015) corresponds to Equation (F-4) (both estimated as linear probability models) when we take \( \hat{H} \) and \( \hat{Y} \) to partition the shift at \{180, 360, 420, 480, 540, 600, 660, 720, 780\} minutes and \{100, 150, 200, 225, 250, 275, 300, 350, 400\} dollars, respectively.

We consider the following stopping rules in which decisions do not depend on earnings:

**Simulation 1:** End the shift with certainty at the end of a trip if hours exceeds 9.5, and stop with independent probability 0.05 at the end of any given trip that ends before 9.5 hours.

**Simulation 2:** Driver \( i \) ends the shift with certainty at the end of a trip if hours exceeds a driver-specific level of hours \( \bar{H}_i \), and stops with independent probability 0.05 at the end of any given trip that ends before \( \bar{H}_i \) hours, where we define \( \bar{H}_i \) as one less than the mean hours across all of driver \( i \)'s shifts in the data.

Appendix Table 2 reports the estimated income effects in both simulations from
Equation (TT), from Equations (F-1) to (F-4) estimated as linear probability model, and from Equations (F-1) to (F-3) estimated as a probit model. We refer to the specification in Farber (2005) as F-3a, and the specification in Farber (2015) as F-3b. Equation (TT) produces the expected result in both simulations that we cannot reject the null hypothesis that the income-related coefficients are jointly zero. Using Equation (F-1), Equation (F-3), or Equation (F-4) leads to the incorrect conclusion in both simulations that income significantly influences the probability of stopping, even though the data are generated precisely so that income has no effect. By controlling flexibly for hours in Equation (F-2), we cannot reject the null hypothesis of no income effects in Simulation 1 since the probability of ending a shift as a function of hours is generated to be identical across drivers; in Simulation 2, however, we incorrectly reject the null hypothesis.

We repeat the exercise 1,000 times for each simulation. While the parametric specifications overwhelmingly produce false positives by incorrectly rejecting the null hypothesis of no income effects, Equation (TT) rejects this null hypothesis at the $x$ percent significance level about in about $x$ percent of simulations. Appendix Figure 2 shows this by plotting the distribution of $p$-values from the non-parametric specification.

Next, we extend Equation (TT) to allow for the probability of stopping to depend on the timing of earnings, analogous to Equation (2):

$$
\text{Pr}(d_{i,n,t} = 1) = \sum_j \left[ (\alpha_j h_{i,n,t} + \gamma_{j,k} \sum_k y_{i,n,t,k} + \mu_{i,j}) \mathbf{1}_{\{h_{i,n,t} \in H_j\}} \right] + \epsilon_{i,n,t} \quad \text{(TT*)}
$$

We simulate a stopping rule in which income does affect stopping decisions, but the timing of income is irrelevant (i.e., money is fungible).

**Simulation 3:** End the shift with certainty at the end of a trip if hours exceeds 9.5, and stop with independent probability $0.05 \cdot y_{i,n,t}$ at the end of any given trip that ends before 9.5 hours, where $y_{i,n,t}$ denotes cumulative daily earnings.

Appendix Table 3 reports the estimated effects of earnings in each hour from Equation (TT). An $F$-test rejects the hypothesis that the income-related coefficients are jointly zero (i.e., $\gamma_{j,k} = 0$ for all $j,k$) but fails to reject the hypothesis that the timing of income is irrelevant (i.e., $\gamma_{j,k_1} = \gamma_{j,k_2}$ for all $j,k_1,k_2$).
D Robustness

D.1 Effort

Section 3.4 discusses IV estimates for the income effect, instrumenting for earnings with the cumulative distance between pick-up and drop-off locations (in GPS distance or odometer miles).

As an alternative estimation strategy, we instrument for cumulative daily earnings using cumulative tips from credit-card transactions. Appendix Figures 3 and 4 reproduce Figures 6 and 7, respectively, using the three IV strategies and find consistent income effects throughout the shift as well as the timing pattern across all specifications. The results in Appendix Figures 3d and 4d show smaller point estimates of the income effect, consistent with our interpretation of the effect of cumulative earnings on the probability of stopping as representing an income effect.

Appendix Figure 5 restricts the analysis to trips that stay within Manhattan and demonstrates that the income effect persists. To understand the effect of cumulative earnings on stopping, this restriction serves several purposes. First, even though the baseline estimates contain an extensive set of controls for location (195 NTA fixed effects and an indicator for being in the zip code where the cab must be returned interacted with hour of the day), the estimates from this subsample ensure that income effects do not only appear when drivers end a trip in one of the outer boroughs (e.g., near the garage where they return the cab or their home). Second, given the dense streets of Manhattan, the variation in earnings due to distance based on differences in driving speed plausibly arises due to traffic conditions unrelated to the driver’s decisions to exert additional effort.

D.2 Learning

Within-day learning

The results in Figures 7 and 8 suggest that drivers react differently to money earned in different hours of the shift, which we interpret as a violation of fungibility. Differences in behavior due to timing of payment could also result from learning. To explain the timing pattern with a learning story, we would have to assume that drivers tend to ignore recent experiences in the market and instead rely on earnings earlier in the shift to predict future opportunities. The data do not support the view that more recent market conditions are less relevant for predicting future market conditions.
(see Figure 5). Instead, a more plausible learning effect would bias the results away from finding stronger effects on stopping in response to more recent earnings. Insofar as within-day learning influences behavior, the estimated violation of fungibility understates the true effect.

**Across-day learning**

In the model from Appendix C.1, earnings on one day convey no information about earnings on another day, so intertemporal optimization is equivalent to maximization of a static one-period objective function. If higher earnings correlate with plentiful opportunities on the next day, then a driver may decide to work less on one day to conserve energy to work more on the next day. The insignificant autocorrelation in the transitory component to daily wages in Appendix Figure 6 suggests little scope for this type of intertemporal substitution to drive the relationship between earnings and quitting.

In a related setting, Agarwal et al. (2015) find that daily income distributions for Singaporean cabdrivers are independent of income shocks in the previous days, which suggests a limited role for intertemporal substitution.

**D.3 Option value**

The trip-by-trip stopping model relies on the assumption that the option value of continuing to drive is sufficiently small (or that drivers ignore option value). The pattern in Appendix Figure 1 suggests that this assumption may not be reasonable for night-weekend shifts, when wages rise substantially and predictably over time, but may be quite reasonable for night-weekday shifts, when the typical wage profile is nonincreasing. Appendix Figure 7 plots the percent change in the probability of stopping estimated on four separate groups of shifts (day-weekday shifts, day-weekend shifts, night-weekday shifts, and night-weekend shifts). We see a significant negative effect of earnings on quitting for night-weekend shifts, for which the assumptions of the stopping model likely do not apply, and significant positive income effects for day shifts and night-weekday shifts, for which the assumptions of the stopping model likely do apply.

---

47 We compute an autocorrelation using market wages, as driver-specific wages are endogenous to their stopping decisions.
D.4 Experience

To investigate the effects of experience, Appendix Table 4 restricts the sample to drivers with over 100 shifts and separates each driver’s shifts into deciles based on the date. The first row corresponds to the first 10 percent of each driver’s shifts, while the last row corresponds to the last 10 percent of each driver’s shifts. The estimates show consistent magnitudes of the income effects as well as violations of fungibility across all levels of experience. For comparison, Haggag et al. (2017) document significant learning among cabdrivers in a relatively short time horizon, with productivity differences between new and experienced drivers vanishing after 17 to 62 shifts.

D.5 Measurement Error

Observability of shift ending

Our empirical approach reveals a decrease in labor supply under the assumption that all shifts end as soon as the driver drops off the last passenger. However, 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. The conclusion that drivers respond to higher cumulative earnings with a reduction in labor supply might be overstated if drivers spend relatively more time searching before quitting in high-income shifts.

Explaining the patterns in our data by the fact that drivers may spend unrecorded amounts of time searching before quitting would require that finding a passenger is more difficult at the end of a shift in which the driver earns more. As noted in the discussion of unobserved effort in Section 3.4, however, a high-income shift is more likely to be one in which the driver generally spends less time searching for passengers. The negative correlation between the share of working hours spent searching for passengers and total earnings in a shift provides suggestive evidence that drivers are unlikely to spend relatively more time searching for a passenger before ending a shift when earnings are high. As an alternative measure of the difficulty of searching at the end of a shift, we use the amount of time that the driver spent searching for the last passenger. Indeed we find a similar pattern: drivers spend an average of 11.2 minutes searching for their last passenger among shifts in the bottom decile of
earnings, compared with only 10.1 minutes among shifts in the top decile.\textsuperscript{48}

The evidence suggests that the income effect does not emerge from the fact that our dataset does not report the amount of time that a driver spends working at the end of a shift. If anything, drivers may spend relatively more minutes working after dropping off the last passenger on a low-wage shift, which would imply that the reduction in labor supply that we observe in response to higher cumulative earnings underestimates the true income effect.\textsuperscript{49}

\section*{Taking breaks}

Appendix Figure 8 classifies breaks as long periods of time without a passenger and presents estimates of the stopping model from Equation (2) with additional controls for minutes spent on break. \textit{Farber (2005)} uses the following thresholds to classify waiting times as breaks: 30 minutes between Manhattan fares; 60 minutes between non-airport, non-Manhattan fares; 90 minutes between airport fares. We test the sensitivity of the income effect by uniformly adjusting the thresholds of waiting time for defining breaks by 15 minutes in either direction. Appendix Figure 8 verifies that the results remain unchanged using these definitions of breaks.

Instead of directly controlling for break time as in \textit{Farber (2005)} (e.g., if taking breaks constitutes an outcome of earnings), we can re-estimate the stopping model using breaks as the dependent variable. A decrease in the probability of taking a break in response to additional earnings might lead to concerns that the stopping model incorrectly attributes the effect of hours work hours to income, but the evidence points against this. We find that an additional 10 percent in earnings corresponds to an increase of 0.0072 to 0.0756 percentage points in the probability of taking a break at 8.5 hours. We find an increase in the probability of taking a break at earlier hours of the shift and no significant change in the probability of taking a break at later hours of the shift.

\textsuperscript{48}The result is also similar if we measure the difficulty of finding a passenger after a given trip by computing the number of minutes spent searching averaged across all drivers whose trips end in the same minute.

\textsuperscript{49}The concern that shift ending times are unobservable might be more relevant for elasticity-based analyses of daily labor-supply decisions, since the fact that drivers spend more unrecorded minutes searching for passengers during shifts with lower average wages could bias elasticity estimates in the positive direction.


E Additional results for structural estimation

The model of loss aversion in Section 4.1.2 makes two simplifying assumptions: first, the targets $I^r$ and $H^r$ represent point expectations, and second, utility is piecewise linear in gains and losses. A stochastic reference point would consist of the distribution of earnings and hours for each shift. We approximate this by using quartiles of the distribution of earnings and hours for each driver and day of the week. The results in Appendix Table 5 show a larger coefficient of loss aversion but still an important role for adaptive reference points and a smaller degree of loss aversion over income. Allowing for diminishing sensitivity corresponds to an objective function that exhibits convexity in losses and concavity in gains. We use the power function

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

where we follow Hastings and Shapiro (2013) by calibrating the parameter $\zeta$ to 0.88 (Kahneman and Tversky, 1979). Appendix Table 6 re-estimates the loss-aversion model with diminishing sensitivity. Overall we find that relaxing the simplifying assumptions in the loss-aversion model does not change the conclusions about the importance of adaptive reference points and the smaller degree of loss aversion over income.
Figure 1: Supply of cabs throughout the day

Note: The figure depicts the average number of cabs that are on the road at any given minute of the day in our cleaned data. The solid line depicts the supply pattern of cabs searching or carrying passengers. The dashed line depicts the supply pattern of cabs with passengers.
Figure 2: Tip distribution by fare

Note: The figure depicts the distribution of tips by fare in the sample for which we observe a tip when tips are between 0 and 20 dollars and fares are between 0 and 60 dollars. The level of darkness represents the density of the points.
Figure 3: 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.
Figure 4: Pattern of wages throughout the day

Note: The figure depicts the average market wage every minute throughout the day from hour 0 to hour 23. The market wage in each minute is the average of the per-minute wages of all drivers working during that minute, where a driver’s per-minute wage is the ratio of the fare (not including tips) to the number of minutes spent searching for or riding with passengers for their current trip. Gray lines are one-standard-deviation bounds over the course of the year 2013.
Figure 5: Autocorrelation of residualized hourly market wage

Note: The figure depicts the autocorrelation of hourly market wages indexed by hour of the calendar year 2013. The hourly market wage is the sum of the minute market wage in each hour, with the minute market wage computed as in Figure 4. The hourly market wage is residualized from a regression on a set of time and weather effects: an interaction between the hour of day and day of week, the week of the year, an indicator for federal holidays, an indicator for whether it rains during that hour, and indicators for high (over 80 degrees Fahrenheit) and low (under 30 degrees Fahrenheit) average hourly temperature. The shaded region denotes a 95-percent confidence band.
Figure 6: 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 at the specified number of hours. The solid lines, corresponding to the scale on the right, depict the marginal effect of an additional 10 percent in earnings on the probability of stopping at various times throughout the shift. Estimates obtain from Equation (1) with controls for location, time, and weather (see Table 2 for details) and fixed effects for 37,460 drivers. The dashed lines represent the 95-percent confidence interval, with standard errors adjusted for clustering at the driver level.
Figure 7: Stopping model estimates: Income effect at 8.5 hours—Timing pattern

Note: The figure depicts the percent change in the probability of ending a shift at 8.5 hours in response to a $10 increase in earnings accumulated at different times in the shift. Estimates obtain from Equation (2) with controls for location, time, and weather (see Table 2 for details) and fixed effects for 36,900 drivers.
Figure 8: Stopping model estimates: Income effect throughout the shift—Timing pattern

Note: The figure depicts the effect of an additional $10 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 (2) with controls for location, time, and weather (see Table 2 for details) and fixed effects for 37,460 drivers.
Figure 9: Stopping model estimates: Income effect at 8.5 hours—Timing pattern by shift start hour

![Graph showing the effect of an additional $10 in earnings accumulated at different times in the shift (vertical axis) on the probability of stopping at 8.5 hours for groups of shifts that start in different hours (horizontal axis). Each square has area proportional to the corresponding percent change in the probability of stopping. Estimates obtain from Equation (2) with controls for location, time, and weather and driver fixed effects (see Table 2 for details).]

Note: The figure depicts the effect of an additional $10 in earnings accumulated at different times in the shift (vertical axis) on the probability of stopping at 8.5 hours for groups of shifts that start in different hours (horizontal axis). Each square has area proportional to the corresponding percent change in the probability of stopping. Estimates obtain from Equation (2) with controls for location, time, and weather and driver fixed effects (see Table 2 for details).
Figure 10: Stopping model estimates: Income effect at 8.5 hours—Data and models

Note: The figure compares the predicted income effects from models of expectations-based loss aversion and salience estimated in Table 5 with the income effects estimated using Equation (2). The confidence interval displays the estimates from Figure 7 of the percent change in the probability of ending a shift at 8.5 hours in response to a $10 increase in earnings accumulated at different times in the shift. The gray square and black diamond represent the predictions of the loss-aversion model and the salience model, respectively.
Figure 11: Stopping model estimates: Income effect at 8.5 hours—Distance from income and hours targets

Note: The figure depicts the percentage-point increase in the probability of ending a shift at 8.5 hours when cumulative earnings is $10 higher. Estimates obtain from Equation (1) extended to include interactions between income, twelve 20-minute intervals based on distance from the hours target, and twelve $10 intervals based on distance from the income target, with controls for location, time, and weather (see Table 2 for details) and fixed effects for 36,900 drivers. A darker shade represents a larger magnitude.
Appendix Figure 1: Pattern of wages: Weekday versus weekend

Note: The figure depicts the average market wage every minute throughout the day, separated into weekdays and weekends. The market wage in each minute is the average of the per-minute wages of all drivers working during that minute, where a driver’s per-minute wage is the ratio of the fare (not including tips) to the number of minutes spent searching for or riding with passengers for their current trip. Weekend is defined as 5 PM Friday through 5 PM Sunday. Weekday is defined as 5 PM Sunday through 5 PM Friday.
Appendix Figure 2: Simulated stopping model: Distribution of $p$ values for non-parametric specification

Note: The figure depicts the results from using Equation (TT) to test the null hypothesis that income has no effect on stopping decisions in Simulation 1 and Simulation 2 repeated 1,000 times each. Simulation 1 denotes a stopping rule in which all drivers end their shifts after exceeding 9.5 hours with some noise in the stopping decision prior to that. Simulation 2 denotes a stopping rule in which all drivers end their shifts after exceeding a driver-specific quantity of hours with some noise in the stopping decision prior to that. The curve represents the cumulative distribution of $p$-values.
Appendix Figure 3: Stopping model IV 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 at the specified number of hours. The solid lines, corresponding to the scale on the right, depict the marginal effect of an additional 10 percent in earnings on the probability of stopping at various times throughout the shift. Estimates obtained from Equation (1) with controls for location, time, and weather (see Table 2 for details) and fixed effects for 37,460 drivers. The dashed lines represent the 95-percent confidence interval, with standard errors adjusted for clustering at the driver level.
Appendix Figure 4: Stopping model IV estimates: Income effect at 8.5 hours—Timing pattern

(a) Baseline
(b) IV: GPS distance
(c) IV: Odometer distance
(d) IV: Tips

Note: The figure depicts the percent change in the probability of ending a shift at 8.5 hours in response to a $10 increase in earnings accumulated at different times in the shift. Estimates obtain from Equation (2) with controls for location, time, and weather (see Table 2 for details) and fixed effects for 36,900 drivers.
Appendix Figure 5: Stopping model IV estimates: Income effect at 8.5 hours—Timing pattern for Manhattan trips

Note: The figure depicts the effect of an additional $10 in earnings accumulated at different times in the shift (vertical axis) on the probability of stopping at 8.5 hours for trips that start and end within Manhattan under various estimation strategies. Estimates obtain from Equation (2) with controls for location, time, and weather and driver fixed effects (see Table 2 for details). Each square has area proportional to the corresponding percent change in the probability of stopping.
Appendix Figure 6: Autocorrelation of residualized daily market wage

Note: The figure depicts the autocorrelation of daily market wages indexed by day of the year in 2013. The daily market wage is the sum of the minute market wage in each calendar day, with the minute market wage computed as in Figure 4. The daily market wage is residualized from a regression on a set of time and weather effects: day of week, week of year, an indicator for federal holidays, an indicator for whether it rains during that day, and indicators for high (over 80 degrees Fahrenheit) and low (under 30 degrees Fahrenheit) average daily temperature. The shaded region denotes a 95-percent confidence band.
Appendix Figure 7: Stopping model estimates: Income effect by shift type

Note: The figure depicts the percent change in the probability of stopping at various times throughout the shift in response to a 10 percent increase in cumulative earnings. Each line represents estimates of Equation (1) (see Figure 6 for details) restricted to the corresponding group of shifts: day-weekday, day-weekend, night-weekday, and night-weekend. Day shifts start between 4 AM and 10 AM, and night shifts start between 2 PM and 8 PM. Weekend shifts consist of night shifts on Friday and Saturday as well as day shifts on Saturday and Sunday.
Appendix Figure 8: Stopping model estimates: Income effect at 8.5 hours—Timing pattern with alternative definitions of breaks

Note: The figure depicts the effect of an additional $10 in earnings accumulated at different times in the shift (vertical axis) on the probability of stopping at 8.5 hours from Figure 7, with controls for break time under various definitions of breaks (horizontal axis). The first column replicates the baseline specification, which controls for minutes spent working, including indicators for the number of minutes in each hour with passengers. The second column uses the following minimum thresholds to classify time spent without a passenger as breaks: 15 minutes between Manhattan fares; 45 minutes between non-airport, non-Manhattan fares; 75 minutes between airport fares. The third column uses the following thresholds: 30 minutes between Manhattan fares; 60 minutes between non-airport, non-Manhattan fares; 90 minutes between airport fares. The fourth column uses the following thresholds: 45 minutes between Manhattan fares; 75 minutes between non-airport, non-Manhattan fares; 105 minutes between airport fares. Each square has area proportional to the corresponding percent change in the probability of stopping.
Table 1: Trip-level summary statistics

<table>
<thead>
<tr>
<th></th>
<th>Median</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ride duration (minutes)</td>
<td>10</td>
<td>12.6</td>
<td>9.2</td>
</tr>
<tr>
<td>Wait duration (minutes)</td>
<td>5</td>
<td>11.3</td>
<td>19.6</td>
</tr>
<tr>
<td>Fare (dollars)</td>
<td>9.5</td>
<td>12.2</td>
<td>9.4</td>
</tr>
<tr>
<td>Tip ratio (percent)</td>
<td>20</td>
<td>19.3</td>
<td>12.3</td>
</tr>
</tbody>
</table>

Note: This table reports summary statistics at the trip level for all 127 million NYC taxi trips in 2013 in the cleaned data (see Appendix A). Ride duration is the number of minutes between pick-up time and drop-off time. Wait duration is the number of minutes between dropping off a passenger and picking up a new passenger. Fare is the amount earned not including tips. Tip ratio is the tip divided by the fare, which is available for the 54% of trips with credit card as the payment type.
Table 2: Stopping model estimates: Income effect at 8.5 hours

<table>
<thead>
<tr>
<th></th>
<th>(1) Baseline</th>
<th>(2) IV: GPS distance</th>
<th>(3) IV: Odometer distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent variable: Indicator for stopping</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Controlling for Hours</td>
<td>0.1442</td>
<td>0.2993</td>
<td>0.4516</td>
</tr>
<tr>
<td></td>
<td>(0.0351)</td>
<td>(0.0459)</td>
<td>(0.0498)</td>
</tr>
<tr>
<td>&amp; Drivers</td>
<td>0.6632</td>
<td>0.9145</td>
<td>0.9974</td>
</tr>
<tr>
<td></td>
<td>(0.0282)</td>
<td>(0.0373)</td>
<td>(0.0354)</td>
</tr>
<tr>
<td>&amp; Location</td>
<td>0.0930</td>
<td>0.1898</td>
<td>0.1789</td>
</tr>
<tr>
<td></td>
<td>(0.0270)</td>
<td>(0.0360)</td>
<td>(0.0334)</td>
</tr>
<tr>
<td>&amp; Time</td>
<td>0.3787</td>
<td>0.3987</td>
<td>0.4141</td>
</tr>
<tr>
<td></td>
<td>(0.0266)</td>
<td>(0.0361)</td>
<td>(0.0330)</td>
</tr>
<tr>
<td>&amp; Weather</td>
<td>0.3798</td>
<td>0.3983</td>
<td>0.4137</td>
</tr>
<tr>
<td></td>
<td>(0.0266)</td>
<td>(0.0361)</td>
<td>(0.0330)</td>
</tr>
</tbody>
</table>

Mean stopping probability: 13.627%
Number of drivers: 36,900

Note: This table reports in each cell an estimate of the percentage-point increase in the probability of ending a shift at 8.5 hours when cumulative earnings is 10 percent higher. Each row specifies an additional set of controls and provides results under three different estimation strategies. Column (1) presents baseline estimates from Equation (1), with $H_j$ partitioning the shift into 10-minute intervals. Columns (2) and (3) instrument for cumulative earnings based on cumulative distance using GPS coordinates and odometer miles, respectively. All specifications control flexibly for minutes spent working, including indicators for the number of minutes in each hour with passengers. Location controls consist of neighborhood fixed effects and an indicator for being in the zip code where the cab must be returned interacted with hour of the day. Time controls include an interaction between the hour of day and day of week, the week of the year, and an indicator for federal holidays. Weather controls consist of the indicators for precipitation, wind speed, and temperature in the minute that a trip ends. Drivers denotes fixed effects for 36,900 unique driver’s license numbers. Standard errors reported in parentheses are adjusted for clustering at the driver level.
Table 3: Stopping model estimates: Income effect at 8.5 hours—Subsamples

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Night weekday</td>
<td>0.3564</td>
<td>0.5421</td>
<td>0.4625</td>
</tr>
<tr>
<td>Medallion owners</td>
<td>(0.0473)</td>
<td>(0.1548)</td>
<td>(0.0805)</td>
</tr>
<tr>
<td>Top decile experience</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income in hour 2</td>
<td>0.0725</td>
<td>-0.1175</td>
<td>-0.0130</td>
</tr>
<tr>
<td></td>
<td>(0.0742)</td>
<td>(0.2351)</td>
<td>(0.1236)</td>
</tr>
<tr>
<td>Income in hour 4</td>
<td>0.0077</td>
<td>0.0282</td>
<td>0.3062</td>
</tr>
<tr>
<td></td>
<td>(0.0717)</td>
<td>(0.2269)</td>
<td>(0.1284)</td>
</tr>
<tr>
<td>Income in hour 6</td>
<td>0.2645</td>
<td>0.2363</td>
<td>0.3309</td>
</tr>
<tr>
<td></td>
<td>(0.0732)</td>
<td>(0.2389)</td>
<td>(0.1267)</td>
</tr>
<tr>
<td>Income in hour 8</td>
<td>0.3270</td>
<td>0.5714</td>
<td>0.5580</td>
</tr>
<tr>
<td></td>
<td>(0.0752)</td>
<td>(0.2246)</td>
<td>(0.1335)</td>
</tr>
</tbody>
</table>

Note: Panel A reports estimates from Equation (1) of the percentage-point increase in the probability of ending a shift at 8.5 hours when cumulative earnings is 10 percent higher. Panel B reports estimates from Equation (2) of the percentage-point change in the probability of ending a shift at 8.5 hours in response to a $10 increase in earnings accumulated at different times in the shift. The columns correspond to different sample restrictions: (1) trips on Friday and Saturday after 5 PM, (2) cabdrivers who operate exactly one cab and no other driver shares that cab, and (3) the latest 10 percent of shifts for drivers with over 100 shifts. The control variables consist of the full set from Table 2. Standard errors reported in parentheses are adjusted for clustering at the driver level.
Table 4: Maximum likelihood estimates: Static reference points

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
<td>Income targeting</td>
<td>Loss aversion</td>
<td>Salience</td>
</tr>
<tr>
<td>Disutility of trip</td>
<td>0.0240</td>
<td>0.0332</td>
<td>0.0450</td>
<td>0.0706</td>
</tr>
<tr>
<td></td>
<td>(0.0021)</td>
<td>(0.0023)</td>
<td>(0.0027)</td>
<td>(0.0040)</td>
</tr>
<tr>
<td>Error term distribution σ</td>
<td>0.0580</td>
<td>0.0644</td>
<td>0.0642</td>
<td>0.0688</td>
</tr>
<tr>
<td></td>
<td>(0.0041)</td>
<td>(0.0049)</td>
<td>(0.0019)</td>
<td>(0.0054)</td>
</tr>
<tr>
<td>Income targeting 1 + α</td>
<td>1.3014</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0031)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loss aversion L</td>
<td></td>
<td>1.3578</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0025)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salience δ</td>
<td></td>
<td></td>
<td></td>
<td>0.3590</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0224)</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-493,669</td>
<td>-474,985</td>
<td>-451,678</td>
<td>-469,585</td>
</tr>
<tr>
<td>Likelihood ratio test: baseline</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: This table presents maximum likelihood estimates of Equation (9) for different specifications of the objective function under the restriction of static reference points. The estimation sample consists of over 1.2 million shifts from 36,900 drivers. The columns correspond to the objective functions in Equations (3) and (5) to (7), and the rows correspond to parameters. The row labeled ‘disutility of trip’ reports a combination of the elasticity parameter $\nu$ and the disutility of effort $\psi$ given by $\psi \frac{H_t^{1+\nu}}{1+\nu}(H_{t+1}^{1+\nu} - H_t^{1+\nu})$. The last row contains the $p$-value from a likelihood ratio test which takes the neoclassical model in column (1) as the null hypothesis.
Table 5: Maximum likelihood estimates: Adaptive reference points

<table>
<thead>
<tr>
<th></th>
<th>(1) Loss aversion</th>
<th>(2) Salience</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disutility of trip</td>
<td>0.0412</td>
<td>0.0694</td>
</tr>
<tr>
<td></td>
<td>(0.0027)</td>
<td>(0.0041)</td>
</tr>
<tr>
<td>Error term distribution $\sigma$</td>
<td>0.0623</td>
<td>0.0702</td>
</tr>
<tr>
<td></td>
<td>(0.0020)</td>
<td>(0.0010)</td>
</tr>
<tr>
<td>Loss aversion $L$</td>
<td>1.3273</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0022)</td>
<td></td>
</tr>
<tr>
<td>Salience $\delta$</td>
<td></td>
<td>0.3290</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0227)</td>
</tr>
<tr>
<td>Adjustment $\theta$</td>
<td>0.9548</td>
<td>0.9672</td>
</tr>
<tr>
<td></td>
<td>(0.0012)</td>
<td>(0.0018)</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-449,346</td>
<td>-465,204</td>
</tr>
<tr>
<td>Likelihood ratio test: $\theta = 1$</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Note: This table presents maximum likelihood estimates of Equation (9) for different specifications of the objective function under adaptive reference points. The estimation sample consists of over 1.2 million shifts from 36,900 drivers. The columns correspond to the objective functions in Equations (3) and (5) to (7), and the rows correspond to parameters. See the note to Table 4 for additional details. The last row contains the $p$-value from a likelihood ratio test which takes the corresponding model with a static reference point as the null hypothesis.
Table 6: Maximum likelihood estimates: Loss aversion over income and hours

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disutility of trip</td>
<td>0.0309</td>
</tr>
<tr>
<td></td>
<td>(0.0030)</td>
</tr>
<tr>
<td>Error term distribution σ</td>
<td>0.0557</td>
</tr>
<tr>
<td></td>
<td>(0.0015)</td>
</tr>
<tr>
<td>Loss aversion over income</td>
<td>1.1037</td>
</tr>
<tr>
<td>$L_I$</td>
<td>(0.0016)</td>
</tr>
<tr>
<td>Loss aversion over hours</td>
<td>1.8628</td>
</tr>
<tr>
<td>$L_H$</td>
<td>(0.0094)</td>
</tr>
<tr>
<td>Adjustment θ</td>
<td>0.8949</td>
</tr>
<tr>
<td></td>
<td>(0.0022)</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-437,900</td>
</tr>
<tr>
<td>Likelihood ratio test: $L_I = L_H$</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Note: This table presents maximum likelihood estimates of Equation (9) for the model of expectations-based loss aversion under adaptive reference points. The estimation sample consists of over 1.2 million shifts from 36,900 drivers. The objective function modifies Equation (6) by allowing for a separate coefficient of loss aversion on each dimension (income and hours). See the note to Table 4 for additional details. The last row contains the $p$-value from a likelihood ratio test which takes the model with a single coefficient of loss aversion as the null hypothesis.
Appendix Table 1: Wage elasticity estimates: Weekday versus weekend

<table>
<thead>
<tr>
<th>Shift</th>
<th>Time</th>
<th>OLS</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day</td>
<td>Weekday</td>
<td>-0.1636</td>
<td>0.2632</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0031)</td>
<td>(0.0074)</td>
</tr>
<tr>
<td>Day</td>
<td>Weekend</td>
<td>-0.1214</td>
<td>0.0076</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0040)</td>
<td>(0.0293)</td>
</tr>
<tr>
<td>Night</td>
<td>Weekday</td>
<td>-0.3546</td>
<td>0.3067</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0027)</td>
<td>(0.0083)</td>
</tr>
<tr>
<td>Night</td>
<td>Weekend</td>
<td>-0.1418</td>
<td>1.3085</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0041)</td>
<td>(0.0260)</td>
</tr>
</tbody>
</table>

Note: Each cell presents elasticity estimates from a regression of log hours on log wages, with time controls (indicators for day of week, week of year, and federal holidays) and driver fixed effects (21,244 for night-shift drivers and 18,569 for day-shift drivers). Day shifts start between 4 AM and 10 AM, and night shifts start between 2 PM and 8 PM. Weekend shifts consist of night shifts on Friday and Saturday as well as day shifts on Saturday and Sunday. For each shift type (day or night), the sample consists of drivers who appear in both the weekday and weekend group. The IV column instruments for wages using the average hourly wage from a non-overlapping sample of 2,108 drivers on the same day. Standard errors reported in parentheses are adjusted for clustering at the driver level.
Appendix Table 2: Simulated stopping model: Decision rule independent of income

<table>
<thead>
<tr>
<th></th>
<th>Simulation 1: stop at 9.5</th>
<th>Simulation 2: stop at $\bar{H}_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect of 20% increase in income</td>
<td>p-value: income coefs. = 0</td>
<td>Effect of 20% increase in income</td>
</tr>
<tr>
<td>Non-parametric model</td>
<td></td>
<td>p-value: income coefs. = 0</td>
</tr>
<tr>
<td>TT</td>
<td>-0.0041 (0.0030)</td>
<td>-0.0058 (0.0043)</td>
</tr>
<tr>
<td>Linear probability model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F-1</td>
<td>-0.0083 (0.0005)</td>
<td>-0.0269 (0.0006)</td>
</tr>
<tr>
<td>F-2</td>
<td>0.0005 (0.0004)</td>
<td>-0.0116 (0.0006)</td>
</tr>
<tr>
<td>F-3a</td>
<td></td>
<td>0.0017</td>
</tr>
<tr>
<td>F-3b</td>
<td>0.0436 (0.0056)</td>
<td>-0.0202 (0.0110)</td>
</tr>
<tr>
<td>F-4</td>
<td>-0.0054 (0.0052)</td>
<td>0.0112 (0.0186)</td>
</tr>
<tr>
<td>Probit model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F-1</td>
<td>-0.0081 (0.0005)</td>
<td>-0.0317 (0.0009)</td>
</tr>
<tr>
<td>F-2</td>
<td>0.0004 (0.0003)</td>
<td>-0.0207 (0.0013)</td>
</tr>
<tr>
<td>F-3a</td>
<td></td>
<td>0.0303</td>
</tr>
<tr>
<td>F-3b</td>
<td>0.0185 (0.0029)</td>
<td>-0.0189 (0.0114)</td>
</tr>
</tbody>
</table>

Note: Each row corresponds to a different regression equation defined in Appendix C.2. The row F-3a corresponds to Equation (F-3) with the partitions $\bar{H}$ and $\bar{Y}$ over hours and income defined as in Farber (2005), and the row F-3b corresponds to Equation (F-3) with the partitions defined as in Farber (2015). Simulation 1 denotes a stopping rule in which all drivers end their shifts after exceeding 9.5 hours with some noise in the stopping decision prior to that. Simulation 2 denotes a stopping rule in which all drivers end their shifts after exceeding a driver-specific quantity of hours with some noise in the stopping decision prior to that. The top panel reports estimates from Equation (TT) of the percentage-point increase in the probability of ending a shift at 8.5 hours when cumulative earnings is 20 percent higher. The middle panel reports results from Equations (F-1) to (F-4) estimated as a linear probability model. The bottom panel reports results from the probit models in Equations (F-1) to (F-4). The income effect (columns 1 and 3) reports the estimated effect of a 20% increase in cumulative daily earnings on the probability of ending a shift after working 8.5 hours and earning $300. The income effect for the model in F-3a mechanically does not predict any effect of income on the probability of stopping after earning $300 because of how the partition is defined. The p-value (columns 2 and 4) presents the result of an F-test (Panels A and B) or $\chi^2$-test (Panel C) of the null hypothesis that the income-related coefficients are jointly zero. The test imposes 1 restriction for Equations (F-1) and (F-2), 9 restrictions for Equation (F-3), 72 restrictions for Equation (F-4), and 59 restrictions for Equation (TT) in Simulation 1, and 78 restrictions for Equation (TT) in Simulation 2.
### Appendix Table 3: Simulated stopping model: Decision rule independent of timing of income

<table>
<thead>
<tr>
<th>TT</th>
<th>Simulation 3: Pr(stop) increases in income</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Effect of 20% increase in income</td>
</tr>
<tr>
<td>Income in hour 1</td>
<td>0.0474 (0.0207)</td>
</tr>
<tr>
<td>Income in hour 2</td>
<td>0.0243 (0.0218)</td>
</tr>
<tr>
<td>Income in hour 3</td>
<td>-0.0012 (0.0223)</td>
</tr>
<tr>
<td>Income in hour 4</td>
<td>0.0102 (0.0225)</td>
</tr>
<tr>
<td>Income in hour 5</td>
<td>0.0215 (0.0227)</td>
</tr>
<tr>
<td>Income in hour 6</td>
<td>0.0241 (0.0212)</td>
</tr>
<tr>
<td>Income in hour 7</td>
<td>0.0285 (0.0208)</td>
</tr>
<tr>
<td>Income in hour 8</td>
<td>0.0157 (0.0228)</td>
</tr>
</tbody>
</table>

*p*-value: income coefs. = 0 0.0000  
*p*-value: Equality of income coefs. 0.8464

Note: Each row reports the estimated percentage-point change in the probability of ending a shift at 8.5 hours in response to a $60 increase in earnings accumulated at different times during the shift from Equation (TT∗). Simulation 3 denotes a stopping rule in which all drivers end their shifts after exceeding 9.5 hours, prior to which the drivers' probability of ending a shift is an increasing function of cumulative daily earnings but does not depend on the timing of those earnings. The penultimate row presents the result of an *F*-test of the null hypothesis that the income-related coefficients are jointly zero. The last row tests the null hypothesis that the $\gamma_{j,k}$ coefficients in Equation (TT∗) are independent of $k$ (for every $j$).
Appendix Table 4: Stopping model estimates: Income effect at 8.5 hours—Within-driver experience

<table>
<thead>
<tr>
<th></th>
<th>(1) Overall</th>
<th>(1) Hour 2</th>
<th>(2) Hour 4</th>
<th>(2) Hour 6</th>
<th>(2) Hour 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–10% shifts</td>
<td>0.2883</td>
<td>0.2984</td>
<td>-0.1003</td>
<td>0.0515</td>
<td>0.1307</td>
</tr>
<tr>
<td></td>
<td>(0.1056)</td>
<td>(0.1558)</td>
<td>(0.1655)</td>
<td>(0.1633)</td>
<td>(0.1723)</td>
</tr>
<tr>
<td>10–20% shifts</td>
<td>0.1874</td>
<td>-0.0933</td>
<td>-0.2133</td>
<td>0.0919</td>
<td>0.5044</td>
</tr>
<tr>
<td></td>
<td>(0.1085)</td>
<td>(0.1516)</td>
<td>(0.1568)</td>
<td>(0.1568)</td>
<td>(0.1622)</td>
</tr>
<tr>
<td>20–30% shifts</td>
<td>0.4576</td>
<td>-0.1167</td>
<td>0.1844</td>
<td>0.3713</td>
<td>0.4914</td>
</tr>
<tr>
<td></td>
<td>(0.1077)</td>
<td>(0.1528)</td>
<td>(0.1590)</td>
<td>(0.1590)</td>
<td>(0.1673)</td>
</tr>
<tr>
<td>30–40% shifts</td>
<td>0.4760</td>
<td>0.0993</td>
<td>-0.1107</td>
<td>0.1451</td>
<td>0.4835</td>
</tr>
<tr>
<td></td>
<td>(0.1085)</td>
<td>(0.1588)</td>
<td>(0.1662)</td>
<td>(0.1597)</td>
<td>(0.1655)</td>
</tr>
<tr>
<td>40–50% shifts</td>
<td>0.3911</td>
<td>0.2189</td>
<td>0.0388</td>
<td>0.1822</td>
<td>0.7382</td>
</tr>
<tr>
<td></td>
<td>(0.1131)</td>
<td>(0.1618)</td>
<td>(0.1719)</td>
<td>(0.1711)</td>
<td>(0.1781)</td>
</tr>
<tr>
<td>50–60% shifts</td>
<td>0.1848</td>
<td>-0.0180</td>
<td>0.0700</td>
<td>0.1722</td>
<td>0.4549</td>
</tr>
<tr>
<td></td>
<td>(0.1152)</td>
<td>(0.1674)</td>
<td>(0.1751)</td>
<td>(0.1758)</td>
<td>(0.1840)</td>
</tr>
<tr>
<td>60–70% shifts</td>
<td>0.3531</td>
<td>0.2247</td>
<td>0.1595</td>
<td>0.2413</td>
<td>0.5915</td>
</tr>
<tr>
<td></td>
<td>(0.1130)</td>
<td>(0.2247)</td>
<td>(0.1595)</td>
<td>(0.2413)</td>
<td>(0.5915)</td>
</tr>
<tr>
<td>70–80% shifts</td>
<td>0.3097</td>
<td>0.0959</td>
<td>0.0022</td>
<td>0.5388</td>
<td>0.4347</td>
</tr>
<tr>
<td></td>
<td>(0.1162)</td>
<td>(0.1647)</td>
<td>(0.1727)</td>
<td>(0.1678)</td>
<td>(0.1765)</td>
</tr>
<tr>
<td>80–90% shifts</td>
<td>0.5603</td>
<td>0.0479</td>
<td>0.3023</td>
<td>0.4168</td>
<td>0.6786</td>
</tr>
<tr>
<td></td>
<td>(0.1150)</td>
<td>(0.1638)</td>
<td>(0.1698)</td>
<td>(0.1689)</td>
<td>(0.1783)</td>
</tr>
<tr>
<td>90–100% shifts</td>
<td>0.4625</td>
<td>-0.0130</td>
<td>0.3062</td>
<td>0.3309</td>
<td>0.5580</td>
</tr>
<tr>
<td></td>
<td>(0.0805)</td>
<td>(0.1236)</td>
<td>(0.1284)</td>
<td>(0.1267)</td>
<td>(0.1335)</td>
</tr>
</tbody>
</table>

Note: Specification (1) reports estimates from Equation (1) of the percentage-point increase in the probability of ending a shift at 8.5 hours when cumulative earnings is 10 percent higher. Specification (2) reports estimates from Equation (2) of the percentage-point change in the probability of ending a shift at 8.5 hours in response to a $10 increase in earnings accumulated at different times in the shift. Each row corresponds to a different level of experience, with the last row denoting shifts with the greatest experience for each driver. The control variables consist of the full set from Table 2. Standard errors reported in parentheses are adjusted for clustering at the driver level.
Appendix Table 5: Maximum likelihood estimates: Loss aversion—Stochastic reference points

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disutility of trip</td>
<td>0.0348</td>
</tr>
<tr>
<td></td>
<td>(0.0022)</td>
</tr>
<tr>
<td>Error term distribution $\sigma$</td>
<td>0.0578</td>
</tr>
<tr>
<td></td>
<td>(0.0031)</td>
</tr>
<tr>
<td>Loss aversion over income $L_I$</td>
<td>1.1164</td>
</tr>
<tr>
<td></td>
<td>(0.0280)</td>
</tr>
<tr>
<td>Loss aversion over hours $L_H$</td>
<td>6.3827</td>
</tr>
<tr>
<td></td>
<td>(0.0978)</td>
</tr>
<tr>
<td>Adjustment $\theta$</td>
<td>0.9060</td>
</tr>
<tr>
<td></td>
<td>(0.0141)</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-350,205</td>
</tr>
<tr>
<td>Likelihood ratio test $L_I = L_H = 1$</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Likelihood ratio test $\theta = 1$</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Likelihood ratio test $L_I = L_H$</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Note: This table presents maximum likelihood estimates of Equation (9) for the objective function in Equation (6) with stochastic reference points. The estimation sample consists of over 1.2 million shifts from 36,900 drivers. See the note to Table 4 for additional details. The last three rows contain the $p$-value from likelihood ratio tests of the following null hypotheses: (i) the baseline model, (ii) a static reference point, and (iii) a single coefficient of loss aversion.
### Appendix Table 6: Maximum likelihood estimates: Loss aversion—With diminishing sensitivity

<table>
<thead>
<tr>
<th>Estimates</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Disutility of trip</td>
<td>0.0816</td>
</tr>
<tr>
<td>(0.0082)</td>
<td></td>
</tr>
<tr>
<td>Error term distribution $\sigma$</td>
<td>0.0795</td>
</tr>
<tr>
<td>(0.0031)</td>
<td></td>
</tr>
<tr>
<td>Loss aversion over income $L_I$</td>
<td>1.2763</td>
</tr>
<tr>
<td>(0.0170)</td>
<td></td>
</tr>
<tr>
<td>Loss aversion over hours $L_H$</td>
<td>2.6971</td>
</tr>
<tr>
<td>(0.0239)</td>
<td></td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-459,015</td>
</tr>
<tr>
<td>Likelihood ratio test: $L_I = L_H = 1$</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Likelihood ratio test: $L_I = L_H$</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Note: This table presents maximum likelihood estimates of Equation (9) for the objective function in Equation (6) with diminishing sensitivity in the gain-loss function. The estimation sample consists of over 1.2 million shifts from 36,900 drivers. See the note to Table 4 for additional details. The last two rows contain the $p$-value from likelihood ratio tests of the following null hypotheses: (i) the baseline model, and (ii) a single coefficient of loss aversion.