RETHINKING REFERENCE DEPENDENCE:
WAGE DYNAMICS AND OPTIMAL TAXI LABOR SUPPLY

NICHOLAS BUCHHOLZ*, MATTHEW SHUM†, AND HAIQING XU‡

ABSTRACT. Workers with variable earnings and flexible hours offer unique opportunities to evaluate intertemporal labor supply elasticities. Existing static analyses, however, have generated well-known puzzles, suggesting evidence of downward sloping labor supply curves. Using a large sample of shifts of New York City taxicab drivers, we estimate a dynamic optimal stopping model of drivers' work times and quitting decisions. Our model exploits a set of sufficient statistics for equilibrium interactions between supply and demand, allowing us to estimate driver opportunity costs via a single agent problem. Our results demonstrate that several apparent behavioral biases documented in the literature can be reproduced using entirely standard preferences. We use our model to provide new estimates of individual earnings elasticities and show that taxi drivers have similar elasticities to workers in markets where experimental evidence has been obtained. Finally, we use data spanning a 2012 fare change to estimate labor supply elasticities with respect to market prices, accounting for the equilibrium impact of prices on supply and demand. We find market elasticities to be approximately a tenth of the size of individual elasticities, suggesting that existing estimates of the benefits to recent earnings legislation in the taxi and ride-hail industries are overstated.

Keywords: Labor supply elasticity, Optimal stopping, Dynamic discrete choice, Taxi industry

JEL: C14, D91, C41, L91

1. INTRODUCTION

Workers with variable earnings and flexible hours offer unique opportunities to evaluate inter-temporal labor supply elasticities. Existing static analyses, however, have generated well-known puzzles, suggesting evidence of downward sloping labor supply curves. In this paper we use data from New York City taxicab drivers to estimate a dynamic optimal stopping model of drivers' work times and quitting decisions. Our data-driven modeling approach and counterfactual questions allow us avoid the need to estimate equilibrium search and matching frictions. We show that the estimated model is capable of reproducing many patterns associated with behavioral biases despite the fact that it specifies a rational-expectations based framework and neo-classical model of labor supply. We use our model to provide new estimates of individual earnings elasticities and show that taxi drivers have similar elasticities to workers in markets where experimental evidence has been obtained.

Studies using observational data aggregated across multiple industries and spanning several years have encountered challenges in measuring labor supply elasticities because observed wage changes typically do not hold all else equal. An earlier literature studies life-cycle models of labor supply (e.g., Heckman and MaCurdy (1980), MaCurdy (1981); Altonji (1986)) and generally finds small positive elasticities. However, the authors acknowledge important limitations imposed by the available data and the assumptions required for credible estimation

As access to rich micro-data has become more available, economists have turned to flexible-work settings, in which workers have the ability to immediately adjust their work time in response to changes in earnings opportunities. These settings appear to offer a clearer path to estimation substitution effects. Yet, the literature has raised several new questions of methodology and interpretation and has therefore not settled on how to obtain the elasticities of interest. A seminal paper is Camerer, Babcock, Loewenstein and Thaler (1997), which finds evidence for negative income elasticities of labor supply among taxi drivers – that is, drivers appear to work fewer hours when faced with higher wage rates. Such a finding is inconsistent with textbook neoclassical labor supply models and instead congruent with alternative "behavioral" labor supply models, such as income targeting, in which agents with flexible hours set an income target and work until the target is reached. A large follow-up literature using a variety of datasets has both corroborated (Ashenfelter, Doran and Schaller (2010), Duong, Chu and

Yao (2022)), expanded (Farber (2008), Crawford and Meng (2011), Thakral and Tô (2021)) or challenged (Farber (2005, 2015)) these findings.

In contrast to the prior literature that has largely treated drivers' earnings expectations as stationary up to predictable time-of-day effects, we demonstrate multiple and significant forms of *within-shift* non-stationary, motivating a forward-looking model of the choice of work hours within a shift. We model taxicab drivers' labor supply decisions as emerging from an optimal stopping problem: in a stochastically evolving environment, drivers provide rides and decide whether or not to continue working or quit for the day. A driver's stopping rule is determined by comparing his expected future earnings at a given place and time against the opportunity cost of continuing to work. This is an inherently dynamic problem, as drivers' choices are derived by comparing tradeoffs in expected future utility levels. Our model offers a simple way to accommodate the complex patterns of serial correlation of earnings. Specifically, we model three state variables: cumulative earnings, cumulative work time and location. We then pair this state vector with a transition probability matrix that is estimated non-parametrically at a level of granularity sufficient to capture the serial correlation patterns. Estimated parameters capture the opportunity costs of drivers as a function of cumulative work hours, allowing for heterogeneity in preferences across different days of the week, daytime or evening shifts, and different types of licenses.

Together with our modeling approach, our first primary contribution is to provide methodology to estimate the intertemporal substitutability of labor and leisure, or the Frisch elasticity of labor supply, from observational data of dynamically optimizing agents. While our model is consistent with equilibrium models of search, matching and spatial sorting in the taxi industry, our approach and specific research questions do not require treating most aspects of these equilibrium outcomes as endogenous. Instead, we are able to specify the model in a highly tractable and parsimonious way with respect to our counterfactuals of interest, allowing us estimate driver preferences as a single-agent dynamic problem.

A second contribution is that our approach can be used to explain both negative wage elasticities, as recovered in Camerer *et al.* (1997) as well as the fact that drivers appear more likely to quit after earnings shocks late into the shift, a "recency" bias documented in Thakral and Tô (2021). In the case of Camerer *et al.* (1997), in which authors analyze end-of-shift data and

3

regress hours worked on average daily wages, we find that the declining relative productivity of drivers on longer shifts leads to negative wage elasticities as a result of a selection effect; a driver with a longer shift will, on average, end the day at a lower productivity state than a similar driver who ends earlier, generating a correlation between low average earnings per hour and longer work hours. In the case of Thakral and Tô (2021), the authors regress drivers' binary decisions to quit on cumulative earnings interacted with the time of day in which the earnings accrued. Their estimates show that drivers exhibit a higher probability of quitting associated with late-in-shift earnings shocks. We find that the particularly large negative autocorrelation of earnings on long trips, for example trips from the New York boroughs of Manhattan to Queens or Bronx, can induce these effects. Long trips appear in the data as earnings shocks and lead to subsequently low expected future earnings. Viewed through the lens of our dynamic stopping model, a positive current earnings shock imposes a less attractive future path of earnings relative to the time cost of continuing to work, leading to a higher probability of quitting for the day.

We validate our model by showing that it is capable of generating the patterns in the data that have been attributed to behavioral explanations. Our results therefore demonstrate that the apparent non-standard behavior can be rationalized as optimal decision making. These findings highlight an important case in which models of perfect rationality produce results that coincide with models of bounded rationality (Rust, 2019). To the extent that workers in flexible work settings actually express non-standard or "behavioral" preferences in their labor supply, our results imply that this behavior may simply serve as a simple heuristic for a solution to a more complicated model of perfect rationality.[1]

Finally, we use our estimated model to explore the effects of a recent policy to raise taxi fares in New York City. The New York Taxi and Limousine Commission estimates that a January 2023 fare change, on average raising trip prices by 23%, will lead to a 33% increase in driver earnings. This estimate implies that labor supply elasticities with respect to fares are no less than 0.40, and only under a strong assumption that demand is perfectly inelastic. We evaluate this elasticity by using data spanning the last such change, a fare hike that raised average prices by 18% on September 4, 2012. Our methodology allows us to analyze these effects without

---

[1]For example, Farber (2005) discusses that, when informally surveyed, a small percentage of taxi drivers report following an income targeting strategy.

explicitly modeling the interaction of supply and demand. We instead exploit the fact that a complete summary of the equilibrium effects of this fare change can be obtained through our estimates of transition densities before and after the policy took effect. With our model estimates in hand, by simulating driver behavior under alternative earnings expectations (as summarized by these transition densities), we find that the elasticity of labor supply with respect to market prices is approximately 0.05, less than one-tenth of our overall Frisch elasticity estimates. This implies that the benefits to individual drivers of a large price increase will be substantially less than predicted by regulators.

Literature. A growing literature has arisen aiming to estimate labor supply elasticities in markets where labor supply is continuously adjustable. Several of these papers have studied the taxi industry, because taxi drivers are typically able to choose their own hours. Moreover, as automated data collection has been implemented to meet regulatory standards, detailed trip data has become available in some of the largest U.S. taxi markets.

Camerer *et al.* (1997) analyzes hand-collected data from New York City and considers the hours worked during individual driver shifts. The authors conduct a series of regressions of log hours worked on the log of average daily wages and find evidence for negative wage elasticities. The authors argue that negative elasticities are consistent with income-targeting on the part of drivers: for example, a labor supply policy of the form "I will work today until I earn $200." Farber (2005, 2008, 2015) consider static optimal stopping models of labor supply. The first paper develops a stopping rule model which explores similar forces to our model, showing that drivers' stopping is most reliably predicted by hours instead of income. The latter two papers integrate reference-dependent utility, which is the notion that agents' utility is not only a function of income but also reference-points or targets, where the marginal utility of income increases more quickly before the target is met than after it is met. While Farber (2008) finds mixed evidence for the existence of reference-dependence, Farber (2015) uses more comprehensive data and finds stronger evidence that drivers have, on average, upward sloping earnings elasticities. Nevertheless, Farber finds that just under a third of drivers exhibit behavior consistent with negative elasticities. Using data on taxi inspections, Ashenfelter *et al.* (2010) finds that drivers who worked before and after fare hikes tended to work slightly less on average afterwards,

suggesting a small negative earnings elasticity.[2] Our paper highlights new data evidence that drivers' relative earnings productivity tends to decline with hours worked. This fact generates a downward bias in the elasticity estimate in the wage regressions due to a selection effect, one which may reconcile these disparate findings. We further show through our model that the key puzzles are reproducible when simulating data directly from the model.

A newer literature presents evidence that drivers' labor supply behavior is dependent on the time of day in which they earn revenue. Crawford and Meng (2011) specifies and estimates a dynamic model of labor supply incorporating reference-dependence in both income and hours-worked during a shift. The authors find evidence that drivers' types of reference-dependence depends on whether their earnings are high or low early in the shift relative to their long-run average. Thakral and Tô (2021) also takes up the question of whether there is a timing dimension to behavioral biases in drivers stopping decisions, showing that more recent income is a stronger determinant of quitting than income earned earlier in a shift. Our paper proposes an explanation for this behavior by showing that earnings shocks are associated with subsequently negative earnings opportunities. Incorporating this fact into our model, we show that we can reproduce the apparent time-inconsistent behavior. Overall, our paper contributes to this literature by proposing a parsimonious model of dynamic optimization without any behavioral parameters that serves as alternative explanation for drivers' behavior.[3]

We also contribute to a literature on structural models of labor supply and market equilibrium in taxi and ride-hail settings. Our model is closely related to the taxi labor supply model of Frechette, Lizzeri and Salz (2019), in which taxi drivers decide how long to work by weighing the utility of earning revenue against the disutility of working longer. We use our framework

---

[2]There are additional studies outside of the context of taxis which consider related questions: Fehr and Goette (2007) demonstrate positive labor supply elasticities in an experiment providing higher payments to bicycle messengers. Andersen, Brandon, Gneezy and List (2014) also show positive labor supply elasticities in an experiment among market vendors in India, explicitly testing for and rejecting reference-dependence. Oettinger (1999) documents an equilibrium increase in labor effort on high demand days among stadium vendors.

[3]There is also an empirical literature that recovers the elasticities of intertemporal substitution as a part of estimating lifecycle models of labor supply, for example Heckman and MaCurdy (1980), Browning *et al.* (1985), MaCurdy (1981), and Altonji (1986). The data used are highly aggregated and often cross-sectional and cross-industry in nature. These studies generally predict small positive labor supply adjustments as a result of increased wage rates. Within the literature, authors regularly highlight significant data limitations and modeling assumptions: for example one must assume that workers are free to choose their own hours and that wage variation is exogenous, which is unlikely to hold in the analyzed settings. It is also a challenge to separate wealth effects from substitution effects, even when long-run panels are used. Nevertheless, our work corroborates the findings of positive substitution elasticities but at a much finer, intra-daily scale instead of workers' lifecycle.

to study the equilibrium effects of policies that were enacted during the sample period. One important difference is that we do not model endogenous search frictions or strategic entry, and instead we rely on non-parametric estimates of the dynamic path of earnings. This data-driven approach captures the intra-daily dynamics of earnings at more granular level, but at the same time abstracts from modeling the underlying mechanisms of market clearing. In essence we substitute with extra data part of the analysis that would otherwise require an additional equilibrium model of search and matching along with several assumptions to make such a model tractable. Chen, Rossi, Chevalier and Oehlsen (2019) estimate labor supply and the value of flexibility in the setting of Uber drivers. Because Uber drivers are able to supply labor in irregular schedules, often as secondary jobs, the labor supply problem is fundamentally different from that of professional taxi drivers. The authors focus on quantifying driver preferences for the flexibility offered by the platform and the opportunity costs of working at different times of day. Buchholz (2022) considers a model of endogenous spatial equilibrium among taxi drivers. We do not model drivers location choices directly, however these decisions are embedded in our model's transition densities, allowing drivers to account for divergent continuation values associated with distant locations and use these to condition labor supply decisions. This suggests a useful approach to accommodate spatial models built from the framework of Lagos (2000), including many additional settings and applications (e.g. Brancaccio *et al.* 2020; Castillo 2022; Rosaia 2023), when counterfactuals address aggregate moments. Petterson (2022) considers a dynamic labor supply model among taxi drivers to estimate a model of reference dependence. In contrast, our model assumes no reference dependence and derives predictions that are consistent with static evidence of reference dependence.

Finally, we contribute to a broad literature that demonstrates how a variety of static empirical puzzles can be rationalized through dynamic models of firm behavior. Examples of these include the puzzle of procyclical labor productivity, rationalized in part by models of labor hoarding among forward-looking firms (Rotemberg and Summers 1990; Burnside *et al.* 1993; Lagos 2006), the puzzle of firms pricing below cost, rationalized through models of learning by doing or predatory pricing (Benkard 2004; Besanko *et al.* 2014), and the puzzle that incurring costly regulations would benefit firms, rationalized through a model of rising rival entry costs (Ryan, 2012).

In Section 2 we present the data used, document important stylized facts and review the main findings of past literature on taxi driver labor supply. In Section 3 we present a dynamic model of drivers' labor supply decision. In Section 4 we discuss the estimation and identification of our model. Section 5 provides estimation results and revisits the literature in light of them. Section 6 uses the estimated model to conduct counterfactuals that measure labor supply elasticity in the context of both individual and market-wide wage fluctuations. Section 7 concludes.

2. DATA AND EVIDENCE

We start by introducing the dataset and presenting some descriptive results. In 2009, The Taxi and Limousine Commission of New York City (TLC) initiated the Taxi Passenger Enhancement Project, which mandated the use of upgraded metering and information technology in all New York medallion cabs. The technology includes the automated data collection of taxi trip and fare information. This data set represents a complete record of all trips operated by licensed New York medallion taxis. We primarily use TLC data on all medallion cab rides given from July,1 2012 to September 3, 2012, the last day before a fare change. The sample analyzed here consists of 27,830,861 trips. Data include the exact time and date of pickup and drop-offs, trip distance, trip time, fare information and car and driver identifiers. Table 1 provides summary statistics on trips in Panel I and driver shifts in Panel II. In Section 6 we complement this data with two additional months of trips that occur after the fare change.

Recent work has made broad use of this data set (See, e.g. Farber (2015), Haggag *et al.* (2017), Frechette *et al.* (2019), Thakral and Tô (2021), Buchholz (2022)). Earlier research, including work devoted to explicitly measuring labor supply elasticities, employs smaller samples and less reliable taxi trip data. While there is continued debate about model specification and the presence of behavioral biases, the TLC data obviates most lingering worries about sample size and measurement error.

There are several regulatory statutes governing TLC licensed taxis that are relevant for analyzing the labor supply of drivers. The TLC licenses are divided into categories, including yellow taxis, liveries, para-transit and special charter vehicles. In 2012, yellow taxis are by far the highest volume of the license classes, providing around 175 million rides per year among roughly 50,000 drivers and 13,437 cars. In 2013 the TLC began licensing "green cabs" or "boro

8

cabs", which grants authorization to pick up passengers outside of Manhattan. Later, they permitted high-volume for-hire vehicles that include platform-based ride-hail companies such as Uber and Lyft.

Yellow taxi medallions are sub-divided into operational types. The most common is the mini-fleet type, representing about 60% of the total yellow taxis, in which companies own multiple cars with attached medallions and drivers lease the taxi on a daily or weekly basis, paying a fixed leasing fee subject to regulated caps. Daily leases impose strict shift limitations, where day-shift drivers are required to return cars to bases, typically located in Queens, by 4-5pm for the evening shift, and night-shift drivers are required to return the car by 4-5am. The remaining types involve a driver-owned car and a leased medallion, or a driver-owned car and driver-owned medallion, neither of which are subject to daily shift restrictions.

Yellow taxis are required to locate passengers through street hail and cannot schedule rides in advance.[4] During the sample period the fare was fixed by the TLC at $2.50 fixed fee plus $2.00 per mile.

Table 1 contains summary statistics for the full sample of trips in Panel I and driver shifts in Panel II, as well as subsamples by weekday vs. weekend, morning vs. evening shift, and fleet status. Morning (AM) and evening (PM) shifts are defined similarly to Farber (2015): AM shifts between 4am – 10am and PM shifts start between 2pm – 8pm. Fleet status is not directly observed but can be partially inferred from data. We're specifically interested in capturing whether an AM-shift driver is obligated to return the taxi by 5pm, or similarly for a PM-shift driver at 5am.[5] Fleet drivers here are interpreted as daily lease drivers, whereas non fleet drivers include owner-operators with and without leased medallions as well as fleet drivers who operate on longer term leases.[6]

---

[4]Permission to engage in app-based e-hailing of yellow taxes, a limited type of scheduling and search aid, was approved by the TLC in mid-December 2012 and took effect in February 2013.

[5]Though exact shift times are not recorded, we take shifts to be drivers' total span of work without breaks longer than five hours. This definition is adopted by some of the literature using this data set (e.g., Haggag et al. (2017), Frechette et al. (2019)) and close to the six hour definition used in the rest (e.g., Farber (2015), Thakral and Tô (2021))

[6]While our data do not directly record the type of license used by each driver, they do identify the medallion ID separately from the driver ID. We categorize leased vs. owner-operator medallions based on the number of unique drivers observed using each medallion and the probability that a medallion is turned over during the "witching hour" between 4pm – 5pm. See Section A.1 for more detail.

TABLE 1. Trip-level Summary Statistics

Variable	Data Sample	Obs.	10%ile	Mean	90%ile	S.D.
	I. Trip Summary Statistics					
	Overall	27.9M	5.60	12.35	21.50	9.23
	Mon-Fri	18.6M	5.70	12.47	21.82	9.39
	Sat-Sun	9.3M	5.50	12.11	20.90	8.90
Trip Revenue ($)	AM Shift	17.5M	5.40	11.80	20.30	8.98
	PM Shift	7.1M	6.40	13.39	23.29	9.33
	Fleet	14.9M	5.52	12.06	20.78	8.75
	Non-Fleet	13.0M	5.70	12.69	22.68	9.74
	Overall	27.8M	4.00	12.15	22.70	8.61
	Mon-Fri	18.6M	4.00	12.50	23.00	8.64
	Sat-Sun	9.3M	4.00	11.45	21.17	7.89
Trip Minutes	AM Shift	17.5M	4.00	11.96	22.23	8.47
	PM Shift	7.1M	4.52	12.56	23.00	8.36
	Fleet	14.9M	4.00	11.90	22.00	8.20
	Non-Fleet	13.0M	4.00	12.45	23.00	8.76
	II. Shift Summary Statistics					
	Overall	1,287K	150.26	267.25	378.59	89.94
	Mon-Fri	880K	153.32	262.18	358.20	81.90
	Sat-Sun	406K	144.77	278.27	419.33	104.46
Shift Revenue ($)	AM Shift	766K	174.97	270.10	359.52	77.07
	PM Shift	331K	166.57	285.60	410.00	92.59
	Fleet	647K	168.06	277.24	380.76	85.33
	Non-Fleet	640K	138.35	257.21	375.64	93.28
	Overall	1,287K	287.92	497.91	659.58	149.60
	Mon-Fri	880K	304.12	497.69	641.27	140.11
	Sat-Sun	406K	262.85	498.37	689.87	168.35
Shift Minutes	AM Shift	766K	380.07	525.71	649.57	125.30
	PM Shift	331K	303.24	485.97	662.10	139.74
	Fleet	647K	331.28	516.79	658.79	137.13
	Non-Fleet	640K	260.59	478.81	661.03	158.97

This table uses TLC Data from August–September, 2012. Panel I summarizes revenues and total trip times over all trips in the sample, with subsampling on different types of shifts as indicated. Panel II summarizes total cumulative revenues and cumulative work time across each driver-shift. Shift minutes includes both the time spent vacant and the time spent occupied.

Panel I shows that nearly all driving yields similar distributions of trips in terms of fares and duration. However weekdays, afternoon shifts, and non-fleet drivers face slightly higher fares and longer durations. These differences reflect a slightly lower concentration of central
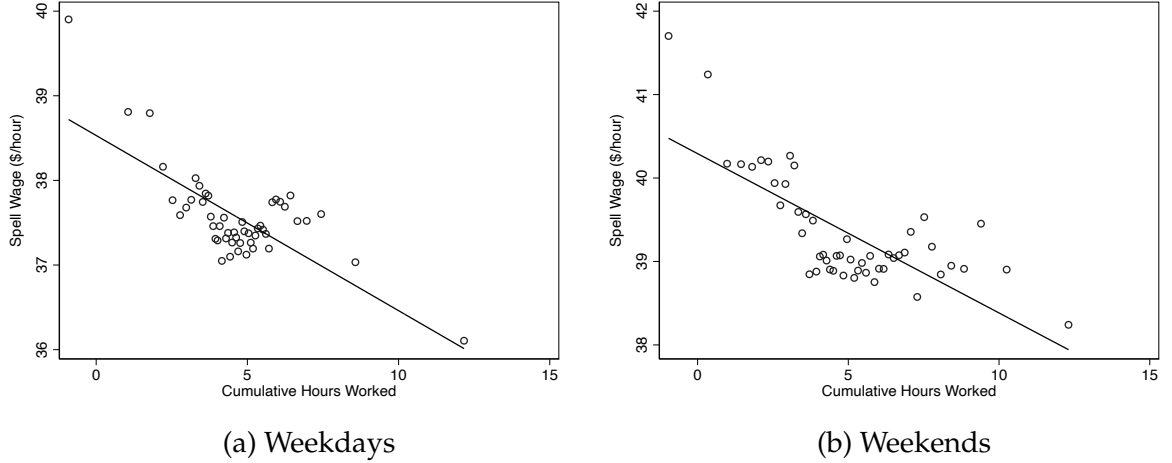
Manhattan trips. Panel II shows that evening shifts earn drivers about 6% more revenue across shifts with about 7% less time. Though owner-operators have essentially free entry across shifts, the fact that their shifts are similarly distributed as lease drivers suggests that evening work may impose higher opportunity cost on drivers despite being more valuable overall. A similar albeit weaker pattern is true for weekend shifts compared to weekdays.

2.1. **The Time Path of Earnings.** In this section we focus on patterns in how drivers accrue earnings over time. While trip prices are strictly regulated, drivers' realized search times are highly uncertain and will generate variability in the time it takes to earn a fixed amount of income. To capture this in a simple way, we define a driver's *spell wage* as the earnings per time (scaled to dollars per hour) from the previous passenger drop-off time to the current passenger drop-off time. Thus the spell wage reflects the total fare earnings from a trip divided by the total search time plus actual travel time for that trip. Note that the weighted average spell wage, where weights are according to the time spent on each spell, is equal to the average realized wage over a given driver's shift, a common moment used in the literature.

We document two key stylized facts about drivers' spell wages. First, they tend to decline with time spent on a shift. Figure 1 shows a relation between spell wages and cumulative hours worked on weekdays (Panel (a)) and weekends (Panel(b)). To summarize the effects over hundreds of thousands of observations each, both panels depict a binscatter plot of spell wages on work hours, where estimates are residualized over month by day-of-week by hour fixed effects. These plots indicate that drivers earn, on average, around $2 less per hour relative to other drivers from the start to the end of a typical shift.

We present more precise measures and additional insights in Table 2, which shows the effects of cumulative time worked on different measures of earnings per hour, including spell wages. Column 1 presents raw correlations without controls for each of four dependent variables. The subsequent two specifications add day of week by hour fixed effects (column 2) and also driver-shift fixed effects (column 3). Taking column 3 as the preferred specification, we see that the productivity of a spell declines over each hour by 23 cents per hour. Thus a typical driver on the ninth hour of the shift would be expected to earn about $2 less per hour than when he started. Panels III, IV and V detail the underlying reasons for this: fares per ride increase by about 49 cents (implying slightly longer trips on average) but not enough to offset increased

11

FIGURE 1. The Time Path of Earnings by Cumulative Hours Worked

(a) Weekdays   (b) Weekends

TLC from January to August 2012. This figure shows two binscatter plots of how the distribution of spell wages evolve with cumulative hours worked. Panel (a) shows the relation between spell wage and cumulative hours worked on weekdays. Panel (b) repeats this for weekends.

search and travel times, which collectively grow by over one and a half minutes per spell per hour worked. Growing search times may indicate longer breaks or less intense search behavior. In section Section A.2 we provide suggestive evidence that drivers do not modulate effort in response to differential earnings opportunities, thus we interpret this decrease in the earnings rate to be exogenous with respect to cumulative earnings. These patterns reflect a declining within-driver productivity: earnings continue to slow down even when market conditions become more favorable for drivers. One important implication of this declining productivity is that, all else equal, driver shifts with longer cumulative work hours will be associated with lower average shift wages.

On top of showing that wages decline on average over the course of a shift, the second key fact is that they also exhibit a negative serial correlation from trip-to-trip on average. This pattern arises because many long and high-earning trips will leave taxi drivers in less desirable locations that require additional search time to find the next passenger.

Table 3 shows the result of an Arellano-Bond AR(1) regression of spell wage on its own lag with driver-shift fixed effects. We see that there is an average effect of lagged spell wage by about -4% on current spell wage, but depending on the destination this may be as high as -74%, as is the case with trips following a Bronx drop-off. Moreover, the share of non-Manhattan drop-offs

TABLE 2. Earnings Per Hour and Search Times by Cumulative Hours

	(1)	(2)	(3)
Effect of cumulative hours worked on:			
Panel I: Wage (cumulative)	-0.106**	-0.317**	-0.252**
	(0.002)	(0.002)	(0.002)
Panel II: Wage (spell)	-0.078**	-0.213**	-0.233**
	(0.006)	(0.007)	(0.009)
Panel III: Earnings/Trip	0.110**	0.085**	0.485**
	(0.003)	(0.004)	(0.007)
Panel IV: Trip Time (sec)	10.81**	7.32**	35.59**
	(0.173)	(0.230)	(0.393)
Panel V: Search Time (sec)	27.95**	32.39**	59.21**
	(0.407)	(0.539)	(0.991)
Dow x Hour	x	✓	✓
Driver-Shift FE	x	x	✓
N	25,689,471	25,689,471	25,689,471

This table documents the effect of the reported variables due to cumulative hours worked.

increases from about 7.7% among drivers on the first four hours of their shift to about 11.6% among drivers who are beyond 8 hours into their shift. The average duration of a spell outside of Manhattan is about 25 minutes, suggesting that drivers who receive a high wage draw in the form of a borough-destination trip face about half an hour of substantial subsequent "wage cuts" in the form of a lower expected spell wage. The negative auto-correlation shown here plays an important role in the dynamic model because it helps explain why drivers often exhibit an increased likelihood of quitting after positive earnings shocks near the end of their shift. Without incorporating forward looking behavior into the labor supply model, these stylized facts would be hard to explain.

In Panel II of Table 3 we show that these patterns coincide with a higher likelihood of taking trips outside of Manhattan and a higher likelihood of quitting after those destinations. For example, during the day shift, highlighted in gray, the share of trips to Brooklyn rises from 1.68% to 3.99% from 6am to 6pm. The likelihood of quitting at a Brooklyn destination also rises from 1% in the early hours to a peak of 35% between 2-4pm, likely reflecting shift quitting in

TABLE 3. Autocorrelation in Spell Wages

	Overall	**Bronx**	Lag Drop-off Borough **Brooklyn**	**Manhattan**	**Queens**
I. AR(1) Regression Coefficients					
Lag Spell Wage	-0.036**	-0.737**	-0.060**	-0.070**	-0.506**
	(0.003)	(0.023)	(0.015)	(0.003)	(0.015)
II. Trip Share (%) / Quit Probability (fraction) in Parentheses					
12AM-2AM	. (0.13)	1.11 (0.27)	14.44 (0.20)	76.61 (0.09)	7.78 (0.29)
2AM-4AM	. (0.22)	1.76 (0.39)	16.30 (0.31)	67.58 (0.16)	14.05 (0.37)
4AM-6PM	. (0.02)	0.84 (0.06)	4.32 (0.07)	82.23 (0.01)	12.08 (0.05)
6AM-8AM	. (0.00)	0.38 (0.01)	1.68 (0.01)	93.59 (0.00)	4.22 (0.01)
8AM-10AM	. (0.01)	0.28 (0.01)	1.63 (0.01)	94.50 (0.00)	3.48 (0.02)
10AM-12PM	. (0.01)	0.28 (0.03)	1.91 (0.04)	93.83 (0.01)	3.85 (0.06)
12PM-2PM	. (0.05)	0.27 (0.15)	2.17 (0.15)	92.66 (0.03)	4.69 (0.24)
2PM-4PM	. (0.19)	0.29 (0.36)	2.54 (0.35)	91.42 (0.17)	5.52 (0.49)
4PM-6PM	. (0.03)	0.37 (0.07)	3.99 (0.06)	91.34 (0.03)	4.17 (0.09)
6PM-8PM	. (0.02)	0.45 (0.06)	5.92 (0.05)	90.35 (0.02)	3.22 (0.05)
8PM-10PM	. (0.03)	0.63 (0.10)	8.66 (0.07)	86.61 (0.02)	4.05 (0.08)
10PM-12AM	. (0.09)	0.84 (0.24)	11.18 (0.17)	82.45 (0.06)	5.49 (0.22)

Panel I of this table reports the results of an Arellano-Bond AR(1) regression to test for the auto-correlation of spell wages. The first specification pools all data. Specifications (2)-(5) subset data according to the last (i.e. lagged) drop-off Borough. Robust standard errors are reported. Asterisks (**) denote significance above the 1% level. Panel 2 shows the percentage of trips with destinations in each boro by time of day. The gray shaded region denotes typical AM shift hours. Percentages do not add up to one because destinations outside these boroughs have very low volume and not reported.

advance of the 5pm shift turnover. In general, the share of destinations outside of Manhattan is increasing towards the end of the day-shift hours and again increases sharply into the evening shift hours as regular shift hours near their end. Quitting is more likely in these destinations as well. We interpret these collective facts as evidence for the impact of negative serial correlation in spell wages influencing drivers labor supply choices.

2.2. **Static Approaches to Recover Wage Elasticities.** We replicate several recent studies employing a static analysis of labor supply. In Section A.4 we show that our specific data sample is able to reproduce the negative wage elasticities by using analogous wage instruments to Camerer *et al.* (1997), the hazard rate patterns documented in Farber (2005) and the recency

effects documented in Thakral and Tô (2021). These effects are reproducible despite using much richer data compared to the first two studies and a different time period from the latter study.

3. MODEL

Taxi drivers drive around the city searching for customers. They provide rides, earning fares each time, and work until deciding to quit for the day. We model the quitting decisions of individual drivers, indexed by i, engaged in daily shifts indexed by j. Time is discrete and each decision period is indexed by t. Market conditions on a given shift are described by a vector x_j, a categorical index for how profitable a specific shift is for the average drivers. x_j summarizes the net effect of market-wide heterogeneity coming from factors like daily weather conditions, demand conditions and road congestion. Given x_j, drivers' payoff-relevant state is described by their cumulative work time h, cumulative earnings r and a location ℓ from which they are searching for passengers. Note that cumulative work time includes both the time searching to find passengers as well as the time spent driving them.

Drivers earn payoffs from cumulative earnings and cumulative time spent working given by $u_i(r, h)$. In the empirical analysis below we make the assumption that money is fungible throughout the day and further normalize the scale of utility to the dollar, so that $\frac{\delta u}{\delta r} = 1$ for all r and h. The function $u_i(r, h)$ is the main primitive object of interest, as it describes how drivers value their time at various levels of work time.

We assume the mean utility of drivers has the following structure, with r equal to cumulative earnings and h equal to cumulative number of hours worked:

$$u_i(r, h) = r + \theta_{1,i} h + \theta_{2,i} h^2 \tag{1}$$

where $\theta_i^c = \{\theta_{1,i}, \theta_{2,i}\}$ is a vector of unknown cost function parameters to be estimated.[7]

After each trip is completed, the driver observes the current state $\{r, h, \ell\}$ and faces a decision. He may quit for the day, in which all fares earned up to that point are kept but all possible future fares for that day are foregone. Alternatively, the driver may choose to keep working for

---

[7]In practice, we will treat drivers as homogenous within eight discrete *types*: AM/PM shifts × Weekday/Weekend shifts × owner-operator/fleet licenses.

one more spell, in which case he draws from a distribution of new fares dr, new work time dh (consisting of both search time and then time-on-trip) and new location ℓ', such that

$$(dr, dh, \ell') \sim \mathcal{F}(r, h, \ell, x). \tag{2}$$

Each draw adds to the stock of cumulative earnings, i.e. $r' = r + dr$ and similarly for work time $h' = h + dh$. The payoff associated with a new fare is the difference between payoffs at the new state from the previous one:

$$\pi_i(dr, dh) = u_i(r + dr, h + dh) - u_i(r, h). \tag{3}$$

At the end of this trip, the driver once again faces the decision to quit or not.

Note that $\mathcal{F}$ plays a critical role in the model: it represents the equilibrium earnings process for the entire shift, summarizing the search process and the efficiency of driver search as a consequence of the thin/thick externalities on both sides of the market. Our counterfactuals of interest operate on $\mathcal{F}$ in a natural and straightforward way, which allows us to circumvent the need to model the underlying search and matching process and avoid the need to take a stance on its form.

Before making a decision, drivers draw an unobserved opportunity cost ϵ_{td} associated with each period t and each choice to quit or not, $d \in \{0, 1\}$. We assume each ϵ_{td} is i.i.d and distributed as Type I Extreme Value with mean zero and scale parameter σ_ϵ. Define $\epsilon_t = \epsilon_{t0} - \epsilon_{t1}$ as the net unobserved opportunity cost of quitting, drawn once per driver per decision period.

Because the payoff-relevant non-stationarity is captured by the cumulative time state h, we hereafter suppress decision-period subscripts t without loss of generality. Driver i's decision problem can then be characterized by a value of quitting or continuing depending his state $\{r, h, \ell\}$, market conditions x, and a draw of ϵ. We can thus write the driver's value function as follows:

$$V_i(r, h, \ell, \epsilon) = \max \left\{ u_i(r, h) + \epsilon, \int_{F(dr, dh, \ell'|r, h, \ell, x), \epsilon'} V_i(r + dr, h + dh, \ell', \epsilon') \right\} \tag{4}$$

Equation 4 summarizes the timing of the decision problem: after dropping a passenger, drivers observe their private signal ϵ about the relative value of continuing work over quitting. If they

decide to continue, they then draw an updated work time and income. We therefore assume they either quit or commit to search until finding another passenger and revisiting the decision after that passenger is dropped off. Since the entire shift occurs within a single day, we assume there is no discounting. However, the draws dh do impact model timing. First, the wage process $\mathcal{F}$ depends on h, so a long spell dh impacts future draws. Second, the driver's cumulative work time also increments by dh, which increases total costs according to the shape of $u(\cdot, \cdot)$.

Existence and Uniqueness of the Optimal Stopping Rule. We show in Section A.9 that under mild conditions a unique solution exists to the optimal stopping rule.

Entry. We abstract from modeling daily entry costs and instead assume drivers' entry timing decisions are fixed. While this is in part supported by regulatory constraints, as daily lease drivers are bound to morning and evening shift timing windows, owner-operators are generally free to begin work at any hour. Because drivers' actual starting time decisions and constraints are not observable, identifying entry costs would require a model of strategic interaction in daily entry timing. We nevertheless believe the most natural way to estimate the substitution elasticity of labor supply, as we do in Section 6.1, is to fix starting times and predict work hours as a function of changes to earnings. This approach offers a direct comparability between our estimates and those obtained in other labor markets. We also want to emphasize that even a simple dynamic model is capable of replicating the puzzles that arise in the labor supply literature. Nevertheless, this assumption limits the kinds of counterfactuals one can consider using our framework.

Competition. Drivers compete for fares with other drivers. This interaction is encoded in the distribution $\mathcal{F}$ of work time, earnings and location draws. All else equal, more drivers in the market tends to shift the mass of this distribution towards longer realizations of h for a given r, as a thicker supply-side leads to increased driver search times.

Modeling $\mathcal{F}$ as an equilibrium outcome poses unique challenges, as it requires coarsening the states and a set of assumptions necessary to construct a mapping between earnings and search frictions (Frechette *et al.*, 2019) and to evaluate the reallocation of vacant drivers due to endogenous location search (Buchholz, 2022). For our questions of interest, however, we can circumvent many of these difficulties by leveraging an in-sample experiment. First, we will ask

how a single driver's work hours would change if only that driver's payoffs increased while holding fixed market supply and demand. For this question we model the counterfactual as acting directly on $\mathcal{F}$. Our second counterfactual asks how a driver's work hours would change if the price of taxi rides changed. To address this question, we estimate the change in $\mathcal{F}$ by analyzing data just before and just after the price of all taxi rides increased. Although any such fare change would influence both the demand and supply sides of the market, the joint effect on driver expectations is entirely summarized in the post-treatment distribution $\mathcal{F}'$. We can therefore use our parameter estimates to recover each driver's own labor supply decision in light of the new $\mathcal{F}'$. We conduct this analysis to highlight the differences between individual and aggregate income elasticities in Section 6.1 and apply this towards understanding the impact of newly implemented fare regulations.

4. EMPIRICAL STRATEGY

In this section, we discuss the estimation of the model presented in Section 3. We adopt a nested fixed-point procedure where for each guess of the parameter vector we first solve for the fixed point of the value functions in Equation 4. Traditionally, non-stationary dynamic models allow for a relatively straight-forward solution value function by backward induction. However, because the time dimension enters our utility specification directly as a state variable, we take the approach of treating value functions as stationary with respect to income and time-worked. The model nevertheless behaves as non-stationary, because time-worked only evolves forward in each decision period until a terminal value. We impose that time-worked beyond 15 hours results in automatic quitting. In practice this assumption is innocuous since more than 99% of driver shifts end before this cutoff and both the data and model parameters will predict that quitting happens much earlier.

4.1. **Discretization of State Variables:** To construct continuation values we finely discretize the state space and generate a set of transition probabilities between each state. We create fifteen uniformly divided bins between the lowest and highest observed values of earnings and time worked within a shift. This leads to income bins from \$2.50 to \$753.33 in fifteen intervals of \$53.63 and time bins from 0 minutes to 1,008 minutes in fifteen intervals of 77.2 minutes.

Together these two grids form a state space of 225 discrete bins over which drivers face value functions and policy functions.

4.2. **Driver Heterogeneity:** To capture the sources of preference heterogeneity consistent with the static empirical analyses discussed in Section A.4, we will focus our analysis of heterogeneity on eight discrete driver types, according to whether a driver's shift is classified as *AM* or *PM* and *Weekday* or *Weekend*, and whether the driver is classified as *Owner-operator* or *Fleet*. Drivers within each group (e.g., AM-Weekday-Fleet) are assumed to have common cost function parameters and common scale parameters on unobserved shocks. Denote the full vector of parameters for type i drivers as $\boldsymbol{\theta}_i = \{\theta_{i1}, \theta_{i2}, \sigma_{\epsilon,i}\}$ and denote the cost-function-specific parameters as $\boldsymbol{\theta}_i^c = \{\theta_{i1}, \theta_{i2}\}$.

4.3. **Market-level Heterogeneity:** Some days are more profitable for drivers than other days, for example weekdays versus weekends, or simply days in which demand is very high or very low. This variation has persistence within a particular shift and will therefore enter into drivers' forward-looking earnings expectations on any given day. We will also use this type of daily market-level variation to create enough richness in our simulations to allow us to replicate the static instrumental variables strategy used in the literature. We incorporate this heterogeneity into our model along three dimensions: *AM/PM* shifts, *Weekday/Weekend* shifts, and we further separate shifts into five types of days, denoted as x_j or the daily earnings quintile x associated with shift j. We compute x_j by first computing the average spell wage of all drivers across each shift and categorizing these into five quintiles, which represent how productive driving is on average. The three above dimensions combine to create $2 \times 2 \times 5$ shift-types indexed by j.

4.4. **Serial Correlation in Earnings:** Our model also accounts for the two forms of within-driver-shift serial correlation discussed in Section 2.1. We describe each of these in turn.

(i) *Declining spell wages*, or the phenomenon in which drivers tend to become less productive as their shift grows longer, are patterns that enter our model through the state transition probability matrix. State transitions determine the relative probabilities of advancing in cumulative earnings and cumulative time conditional on a location ℓ and daily shift-type x_j. Transition probabilities incorporate declining spell wages documented in Section 2.1 because, as the cumulative time state grows, a driver's probability

19

of reaching higher earnings states declines relative to his probability of reaching higher time states.

(ii) *Location effects*, or the chance that future spell wages fall following trips to outer boroughs, also enter drivers' expectations and impact their quitting decisions. Each time drivers in location ℓ draw a new location ℓ' they earn fare, cumulate work time, and then face a new decision at location ℓ'. Drivers face a location-specific value function $V(\cdot, \cdot, \ell', \cdot)$ which reflects that, for $\ell' \notin$ Manhattan, the expected productivity of earnings falls as a large fraction of drivers will not find a passenger until returning to Manhattan.

To incorporate these locations effects, we model outer-borough decisions in the following way.[8] When a driver drops off a passenger outside of Manhattan, he faces the option to end the shift and collect the current net utility from income and hours, or else return to Manhattan to continue search. The value of search in Manhattan is offset by additional travel time to return there and therefore the value is computed as $V(r, h | \ell \notin Manhattan) = V(r, h + z_\ell | \ell = Manhattan)$ where z_ℓ is the expected travel time to return to Manhattan for search. We estimate z_ℓ for each outer borough by regressing drivers' total searching times on the previous drop-off borough, relative to Manhattan, net of a full set of controls for hour-of-day, day-of-week, shift, owner-operator, and driver-shift fixed effects. See Section A.7 for more details.

4.5. **Estimation.** Our model parameters reflect the tradeoff between earnings and time as revealed by drivers' quitting decisions conditional on observed states. To estimate the model parameters governing this tradeoff we specify a log-likelihood function as follows, among drivers $i \in \mathcal{I}(j)$ of shift-type j and trips t:

$$LL(\theta_j) = \sum_{i \in \mathcal{I}(j)} \sum_{t=1}^{T} \left\{ y_{it} \cdot \left[\ln P(y_{it} = 1 | r_{it}, h_{it}, \ell_{it}, x_i; \boldsymbol{\theta}_j) + \sum_{s=1}^{t_i - 1} \ln P(y_{i,s} = 0 | r_{i,s}, h_{i,s}, \ell_{i,s}, x_i; \boldsymbol{\theta}_j) \right] \right\}$$

(5)

---

[8]Due to dimensionality concerns with adding more locations we treat the space of locations fairly coarsely such that there is one for each of New York's five borough's and one for Newark Airport outside of New York.

The quitting probability $P(y_{it} = 1|r_{it}, h_{it}, \ell_{it}, x_i; \boldsymbol{\theta}_j)$ is obtained via the equilibrium value functions as specified in Equation 4.[9] The functional form assumption on opportunity cost shocks ϵ_{it} allow us to represent these probabilities as:

$$P(y_{it} = 1|r_{it}, h_{it}, \ell_{it}, x_i; \boldsymbol{\theta}_j) = \frac{exp(u(r_{it}, h_{it}|\boldsymbol{\theta}_j^c)/\sigma_{\epsilon,j})}{exp(u(r_{it}, h_{it}|\boldsymbol{\theta}_j^c)/\sigma_{\epsilon,j})) + exp(\int_{F,\epsilon} V_i(r_{it} + dr, h_{it} + dh, \ell', x_i|\boldsymbol{\theta}_j^c)/\sigma_{\epsilon})} \tag{6}$$

Our estimator maximizes Equation 5 across eight distinct observable types of driver-shifts, or $j \in \{1, ..., 8\}$, as reported in Section 5.

5. RESULTS

In this section we discuss the empirical results for the dynamic optimal stopping model of taxicab driver labor supply, which we presented in Section 3. The estimates of the driver cost parameters are reported in Table 4 Panel I. We produce estimates on eight samples, dividing drivers' into owner-operator or fleet and their shifts into day or evening and weekday or weekend. As discussed in Section 2, these are natural divisions across which opportunity costs should differ. For estimation, we normalized the earnings parameter (in dollars) to one, allowing us to interpret cost functions in dollar terms.

We use the mean cost parameters θ_1 and θ_2 to compute drivers' marginal cost of time at quitting, presented in Table 4 Panel II. This value is computed as the mean of the derivative in total time cost with respect to hours worked (i.e., $mc_i(h) = \theta_{i1} + 2\theta_{i2}h$) where h is the final hour of each driver's shift. For example, the first column shows that owner-operators during daytime weekday shifts have an average cost of time at quitting of \$35.50 per hour. We contrast with Panel III, which shows the average shift earnings in that hour to be \$36.94. The similarity of these numbers serve as a measure of fit; the average end-of-shift earnings appear to roughly coincide with estimated time costs. Since time costs are rising rapidly by 7–8 hours, these estimates appear to broadly rationalize drivers' quitting behavior around the observed times. Finally,

---

[9]As a finite-time model, we are able to use non-standard distributions of unobservables, potentially correlated with state variables or serially-correlated. Our estimation procedure would instead need to incorporate simulation methods (Pakes, 1986). While there are some attractive reasons to build these features into the model, our primary goal is to demonstrate that even a simple dynamic model is sufficient to reproduce the labor-supply puzzles document in prior literature.

Table 4 Panel III also displays average shift durations by group. Average shift durations across groups are between 7h 40m and 9h40m.
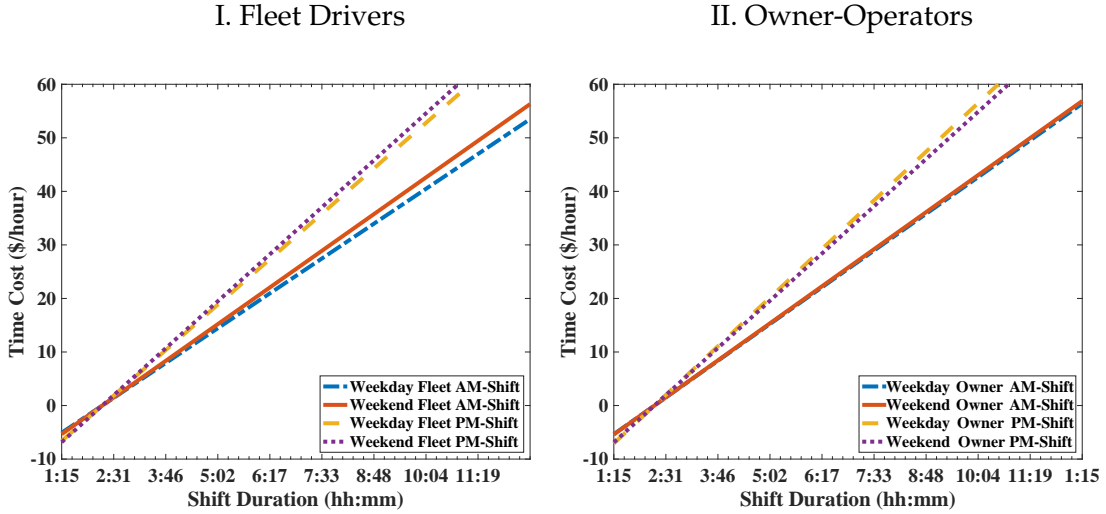
TABLE 4. Model Estimates

Driver-Type:	Owner-Operated		Fleet		Owner-Operated		Fleet	
	AM		AM		PM		PM	
Shift-Type:	Mon-Fri	Sat-Sun	Mon-Fri	Sat-Sun	Mon-Fri	Sat-Sun	Mon-Fri	Sat-Sun
I. Estimates								
σ_ϵ^2	44.03	52.92	46.21	57.11	55.31	76.83	58.95	69.09
	(2.299)	(7.016)	(3.259)	(5.429)	(2.927)	(6.351)	(1.057)	(5.225)
θ_1	60.70	71.16	45.69	64.52	36.47	59.84	45.94	58.69
	(3.224)	(8.778)	(6.093)	(9.260)	(3.371)	(6.113)	(2.334)	(4.935)
θ_2	-5.99	-6.66	-5.32	-6.44	-5.43	-7.31	-6.49	-7.29
	(0.201)	(0.502)	(0.380)	(0.564)	(0.341)	(0.434)	(0.274)	(0.350)
II. Implied by Estimates								
Last Hour Time Cost ($/hr.)	35.50	38.17	35.14	36.28	42.54	38.59	41.95	38.57
	(1.927)	(1.305)	(2.145)	(0.793)	(1.509)	(1.658)	(1.754)	(2.853)
III. Data Comparison								
Last Hour Earning ($/hr.)	36.94	39.68	37.57	40.07	38.55	40.89	37.80	41.26
Avg. Shift Minutes	472	399	509	456	577	568	550	567

This table shows model estimates by shift, owner-status and weekday/weekend. Panel I shows parameter estimates as well as standard errors. Panel II shows the average marginal time cost of drivers at the time of quitting. This cost is computed as the mean (across drivers in each group) of model estimates of drivers' time costs at the time that driver quits. Panel III displays average cumulative earnings and work time for driver shifts within each group. Standard errors are obtained by resampling entire driver shifts, with replacement, and re-estimating state transitions, parameter estimates (panel I) and associated moments (panel II) within each driver- and shift-type. We conduct the estimation across 200 samples for each group and report standard deviations in parentheses.

While the empirical specifications in Table 4 are simple – payoff functions have are only three parameters – the behavioral implications of the dynamic model are quite rich. In Figure 2 we present drivers' marginal cost of time, or $\delta\pi(r, a)/\delta a$. These marginal cost functions are linear as $\pi(\cdot, \cdot)$ is quadratic in cumulative time-worked. An immediate consequence of these estimates is that individual driver labor supply elasticities are positive at the margin, because as earnings grow, drivers' optimal amount of cumulative work time will increase on average, net of the unobserved cost variation. AM preferences are however more alike than PM preferences; the steeper costs in AM shifts may reflect the influence of the 5pm shift change as well as greater preferences for traditional work hours. Another observation is that cost functions themselves are more homogenous than is apparent in Table 4, this is because θ_1 and θ_2 are somewhat negatively

FIGURE 2. NYC Taxi Driver Marginal Cost of Time Estimates

I. Fleet Drivers   II. Owner-Operators

This figure shows estimates of drivers' marginal cost of time by day/evening and weekday/weekend shifts.

correlated. We see in Panel II that AM and PM weekend shifts are similar in average costs as are AM and PM weekday shifts, than are weekday and weekend shifts across a given time of day. Finally, owner-operators appear to have almost identical preferences to fleet drivers when comparing comparable days and times of day.

Nevertheless, a point of emphasis here is that once taxi drivers' quitting decision are modeled in a dynamic optimal stopping framework, the existing static wage or hazard models make less sense as they largely require drivers to react to past outcomes instead of future ones. However, in the dynamic context we can also see why past outcomes matter: drivers react to past earnings shocks (i.e., observed "high" draws from $\mathcal{F}$) because there is autocorrelation in these shocks, at times positive and other times negative. Drivers react to cumulative time because it moves them up their cost curves as illustrated in Figure 2. Both types of serial correlation lead static models to identify real effects of current time and income, but absent the dynamic model the interpretation is difficult.

5.1. **Generating the Behavioral Results.** In this section we show that data generated from our estimated model produces driver behavior that appears as non-standard or "behavioral" when analyzed in a static model. Our first step is to use our model to simulate data that takes the same form as our actual data set. To do this we simulate a large set of driver shifts conditional

23

on driver types (owner-operator or fleet drivers), shift-types (morning or evening shift periods), and day-types (one of five levels of average daily earnings), where drivers begin at an initial state and take draws from the joint distribution of earnings and time when they choose to keep working. We design this exercise to generate a new data set such that driver-, shift-, and day-types are represented in proportions identical to those found in the original data set. We then conduct a series of regressions analogous to those used in the literature and show that the results indicate apparent downward-sloping labor supply curves. In Section A.8 we provide a detailed description of the simulation exercise.

In Table 5 we report estimated wage regressions next to analogous regressions from the original data set. Both data and estimates underlying the simulation come from a two-month period of July 1, 2012 to August 31, 2012. The first two columns display instrumented wage regressions comparable to those in Camerer *et al.* (1997).[10] We report results using two separate instruments representing the distribution of wages among drivers on a given day. *IV 1* denotes the Camerer *et al.* (1997) wage instrument, or the 25th, 50th and 75th percentiles of the average shift wages of other drivers on the same day. *IV 2* denotes an alternative instrument as used in Farber (2015) equal to the average of the daily shift wage rates across all drivers. Under all specifications we find significant negative coefficients on log wage, indicating that shifts in which drivers who earn higher wages are correlated with shifts in which drivers work less time.[11] Columns (3)-(4) replicate identical specifications by directly using the data. While we replicated the specifications of Camerer *et al.* (1997) in Table 10, here we demonstrate that the negative coefficients still obtain when we compare them to our simulated regressions. While our model inherently abstracts from the richness of the decisions taken by actual drivers on the street, our elasticity estimates are slightly more negative than those produced using the actual data. This suggests that, despite its simplicity, our model fully captures the "behavioral" aspects of driver behavior as documented in the early literature.

Despite the apparent negative wage elasticities, we know that our model is consistent with standard preferences over money by construction. To see why we obtain negative coefficients

---

[10]The exception is that our estimates are not conditioned on weather and temperature variables so we omit these covariates across the table.

[11]In columns (1)-(2), which use data simulated from our estimated model, a "day" is simply the combination of *day types d* and weekday or weekend. In columns (3)-(4), which use actual data, a day is defined as the calendar day in which the shift began. Within a day, average shift wages are separated between am and pm shift workers.

TABLE 5. Simulated Wage Regressions Comparison

	(1) Simulated (IV 1)	(2) Simulated (IV 2)	(3) Data (IV 1)	(4) Data (IV 2)
Log Wage	-0.420**	-0.399**	-0.190**	-0.158**
	(0.009)	(0.010)	(0.017)	(0.018)
Weekday	0.138**	0.138**	0.047**	0.046**
	(0.004)	(0.004)	(0.003)	(0.003)
PM Shift	0.231**	0.231**	0.013**	0.011**
	(0.004)	(0.004)	(0.003)	(0.003)
Owner Operator	0.275**	0.274**	-0.055**	-0.056**
	(0.004)	(0.004)	(0.002)	(0.002)
N	51,468	51,468	382,241	382,241

TLC Data from July-August, 2012. Panels (1)-(2) use simulated driver shift data. Data record the final cumulative hours and average wage earned as of the last trip of each simulated driver-shift. IV 1 denotes wage instruments are the 25th, 50th and 75th percentile across all driver wages in each day type, weekday/weekend and am/pm shift. IV 2 denotes wage instruments is the mean of average hourly driver wages across all drivers in each day type, weekday/weekend and am/pm shift. Panels (3)-(4) use TLC data and report the same regressions, where IV 1 denotes the wage instruments are the quartiles of hourly driver wage each day and IV 2 denotes wage instruments that are the average hourly driver wages across all drivers. Standard Errors clustered at the driver-shift level. Asterisks (**) indicate significance at or above the 1% level.

on wage in hours worked, and positive coefficients on earnings in the probability of quitting, we turn back to Section 2.1. There we document how spell wages of a given driver tend to decline relative to other drivers the longer his shift grows. Thus, by evaluating the average wage and the total time worked, we find that longer shifts are associated with a lower average earnings-per-hour compared to drivers who worked shorter shifts as a result of the pattern of declining spell wages.

This pattern holds despite an instrument for wages that generates exogenous variation in the average earnings per hour. This result indicates that a selection effect of lower average wages among drivers with long hours worked dominates the extent to which longer hours are worked due to higher earnings. We show in Section 6.1 that the dynamic model can disentangle these two channels in order to estimate the actual impact of a persistent increase in earnings on total work hours.

Next, we turn to the case of time-inconsistent preferences highlighted in Thakral and Tô (2021). This result again suggests a non-standard, downward-sloping labor supply curve in

25

certain periods of time close to the end of the work day. We now show that the non-stationarity of spell wages can, as before, account for these results. In addition to the fact that relative spell wages tend to decline with hours worked, Section 2.1 also documents a negative serial correlation in earnings per hour. We find this effect to be present, on average, across all trips, but especially so on longer trips to the outer boroughs of New York. Such long trips are also positively correlated with hours worked, as they occur more during end-of-shift hours for both day and evening shifts.

Negative serial correlation implies that large positive earnings shocks will on average be followed by a negative earnings shock. For example, following a trip with spell wage equal to the daily average earnings-per-hour, or about $37, we predict the next period's earnings to be approximately $1.37 per hour lower (or a decline of 3.7%). However, for a trip with a destination in the Bronx, we expect the subsequent spell wage to be approximately $27.27 lower than average (or a decline of 73.7%). As a consequence, earnings shocks – particularly those involving travel to distant regions – represent substantial negative shocks to future earnings and induce a higher likelihood of quitting.

By incorporating negative serial correlation into the dynamic model, our simulations produce data that align with the time-inconsistent preferences found in the literature. To show this, we estimate the following equation:

$$Pr(d_{int} = 1) = \sum_{\ell} \beta^{\ell}(h_{int}) r_{int}^{\ell} + X_{int}\gamma + \epsilon_{int} \tag{7}$$

Here d is the binary decision for driver i to quit or not on shift n at time t. As in Thakral and Tô (2021) we allow drivers' decisions to depend on cumulative earnings r which are earned in hour ℓ of the shift. X includes controls for hour and shift.[12] The specification is then estimated on five different groups: drivers that quit with cumulative hours equal to 7, 8, 9, 10 and 11. In each group, we collect only the observations where drivers quit in the middle 20 minutes of the cumulative hour (i.e., hour 7 quitters are drivers that quit between 7:20-7:40 minutes into their

---

[12]Note we do not include the full set of weather controls and driver fixed effects because these are not estimated separately in our model. However, these controls are not pivotal to the outcomes documented in Thakral and Tô (2021).

shift), and we control for income earned in the previous hours starting at the fourth cumulative hour, not including the final hour.

Table 6 shows the results of this exercise. They reveal that (i) the final hour or two of income is most significant whereas other hours mostly result in insignificant estimates and (ii) for later quit times, the coefficient on income is both positive and significant. Although our simulated sample is much smaller and, by construction, less rich than the TLC data itself, these results suggest we can replicate patterns consistent with adaptive reference dependence or late-in-day income targeting without any explicit modeling of these phenomena, solely through incorporating forward looking decisions into an otherwise simple preference specification.

TABLE 6. Simulated Adaptive Hazard Rates

	(1) Quit Hour 7	(2) Quit Hour 8	(3) Quit Hour 9	(4) Quit Hour 10	(5) Quit Hour 11
Hour 4 income	0.035**	-0.001	0.006	0.010	0.015*
	(0.003)	(0.004)	(0.004)	(0.005)	(0.006)
Hour 5 income	-0.016**	0.021**	0.008	-0.006	-0.020*
	(0.004)	(0.005)	(0.005)	(0.007)	(0.008)
Hour 6 income	0.011**	-0.017**	-0.003	-0.002	0.011
	(0.003)	(0.004)	(0.004)	(0.002)	(0.008)
Hour 7 income		0.007**	-0.022**	-0.001	-0.017*
		(0.002)	(0.003)	(0.005)	(0.007)
Hour 8 income			0.049**	-0.031**	0.002
			(0.002)	(0.004)	(0.007)
Hour 9 income				0.059**	-0.025**
				(0.003)	(0.005)
Hour 10 income					0.048**
					(.003)
N	16,364	11,179	8,370	6,893	5,306

This table replicates the results of Figure 4 of Thakral and Tô (2021). Data are simulated from the dynamic choice model. Asterisks (*) and (**) indicate significance at or above the 5% and 1% level.

5.2. **Discussion: Are Taxi Drivers Dynamic Programmers?** On one hand, the results in Table 5 suggest that drivers may, on average, engage in labor supply behavior that is consistent with positive Frisch elasticities. On the other hand, we impose on drivers a complicated optimization problem necessary to make decisions. In fact, we do not necessarily believe that our results reflect

the direct behavioral process of drivers or even reject the notion of income targets posited in the literature. We do, however, aim to make the point that drivers behavior, even under income targeting or other forms of reference dependent preferences as posited in a myopic or short-run context, may indeed be consistent with the solution to a complex dynamic optimization problem. Our results suggest that approaches like income targeting may constitute "simple" heuristics that drivers have learned, for example through months or years of experience, to maximize their own personal stream of payoffs.

6. ANALYSIS OF LABOR SUPPLY ELASTICITY AND THE EFFECT OF RISING FARES

In this section we examine labor supply elasticity from both the individual and aggregate perspectives. Using these results, we can then estimate the effects of the proposed taxi wage policy on labor supply in the setting of NYC taxi drivers

6.1. **Estimating Individual Labor Supply Elasticity.** Our estimated model allow us to further recover an estimate of the labor supply elasticity of taxi drivers with respect to wage rates. Given the day-to-day nature of earnings variation and driver labor supply choices, it is natural to interpret these as intensive-margin intertemporal substitution (Frisch) elasticities. While these estimates are the target of the literature reviewed in Section A.4, we have shown that existing approaches are flawed when it comes to identifying them. We estimate these wage elasticities by constructing counterfactual wage rate increases and simulating driver shifts and expected work hours with and without the increases. For example, we consider a 10% increase in all earnings available to a *single driver* across the joint distribution of earnings and time draws. This would represent a 10% increase in the measurement of wages preserving the stochasticity of wages as well as the negative auto-correlation discussed in Section 2.1. This counterfactual assumes that demand as well as the supply of all rivals is held fixed. In other words, we assume that when a single driver faces the earnings increase, this driver's impact on demand and any spillover effects to other drivers are negligible.

Table 7 displays the estimated elasticities of hour worked with respect to earnings rates. They show that for a range of earnings increases from 5–25%, the estimated mean hours worked, across all driver types, increases by 2–5 hours, implying individual elasticities between 0.59–0.91. Due to the convexity of costs, elasticity is positive and falling in additional earnings. Table

28

TABLE 7. Individual Labor Supply Elasticity

Earnings Change	Hours Worked			Implied Elasticity	
	P25	Mean	P75	Mean	Std. Err.
Baseline	4.80	7.55	7.86	.	.
5% increase	5.02	7.89	8.67	0.905	0.090
10% increase	5.29	8.06	8.79	0.676	0.048
18% increase	5.32	8.35	8.94	0.589	0.026
25% increase	5.27	8.68	9.59	0.599	0.024

This table reports the distribution of estimated work-hours resulting from simulating shifts at the baseline as well as across a series of increases to the earnings from each trip, assuming that demand and all other taxis behavior remains fixed. Implied elasticities are computed at the mean of hours worked. Standard errors are obtained by resampling entire driver shifts, with replacement, and re-estimating model parameters for each driver-shift type and re-simulating data in equal proportion to how each driver-shift is distributed. We conduct the exercise across 200 samples and report standard deviations in parentheses.

5 in Farber (2015) estimates aggregate elasticities to be 0.589. That estimate comes from a regression of log hours on log wage, similar to the approach of Camerer et al. (1997), albeit with a different dataset, specification and sample period. The estimate therefore corresponds to a 1% increase in earnings. That our estimates are higher is not surprising: there is a downward bias in the reduced form approach due to drivers' within-shift declining productivity of earnings.

We refer to these estimates as *individual* labor supply elasticities because they hold search times fixed, implying that the labor supply of other drivers and consumer demand are both held fixed. If a wage increase were implemented on the entire market, other drivers would change their behavior and consequently lead to endogenous changes in search times. In addition, in this exercise we are holding fixed passenger prices: if passenger prices increased, this would decrease demand for taxi rides and again lead to longer search times. Therefore our estimates can be interpreted as the effect of an exogenous subsidy to an individual driver holding all else equal.

Our individual elasticity estimates are comparable to other labor markets where wages do not impact demand or supply. Looking to other comparable settings, out estimates closely match Table 1 of Chetty et al. (2011), which among other statistics provides an extensive meta-analysis of quasi-experimental evidence on intensive-margin labor supply elasticities, highlighting an average inter-temporal substitution elasticity of 0.54.

We implicitly assume that in a given day the earnings process $\mathcal{F}$ is exogenous. One threat to our strategy would be the case that drivers endogenously choose their levels of search effort, perhaps at higher cost, when the earnings profile changes. In Section A.2 we conduct a simple test for this by analyzing drivers who achieve unexpectedly high earnings early in the shift. We find that such high earnings do not predict subsequently higher earnings, implying that drivers who perceive the day to be more profitable than it is do not earn higher profits in later periods. This suggests that endogenous effort is not confounding our analysis of earnings elasticities.

Although individual elasticities are an important quantity of interest, we can also use our data and empirical approach to learn more. The reason for including the specific case of 18% is that in September, 2012 the NY TLC hiked fares by an average of 18%, which affected search times through both passenger demand and equilibrium effects of labor supply. In the next section we use this change in fares to evaluate and compare individual labor supply elasticities with aggregate or equilibrium labor supply elasticities and consider these effects in the context of recent minimum wage legislation applied to taxi and ride-hail workers in New York City.

6.2. **Equilibrium Elasticities and Wage Policy.** On December 19, 2022 the New York Taxi and Limousine Commission implemented the first fare hike in ten years under a proposal known as "Raise For All".[13] Fares increased through a mix of increased base fares and increased surcharges during rush hour trips and trips to airports. The net effect is estimated by the TLC to increase average passenger fares by 23%.[14]

We investigate how much work hours will be impacted by the rising fares and whether the data generated by such a change be informative about drivers' earnings elasticities. While data are not available to study the 2022 event, nor is it likely to be, as the TLC ceased publication of driver and medallion identifiers, we study this question by considering the last time in which fares were changed. On September 4, 2012 the NYC TLC raised base taxi fares across for all drivers and customers. The distance fee increased from $2.00 per-mile to $2.50 per-mile, the JFK airport flat-fee increased from $45 to $52 and the Newark Liberty Airport surcharge increased

---

[13]See details in Hartwell *et al.* (2022).

[14]This estimate compares the existing distribution of routes and multiplies trips the new fares. In reality, and even by ignoring potential supply-side effects, the demand response alone on more affected routes would likely dampen this aggregate estimate.

from $15 to $17.50. Collectively these changes amounted to an 18% increase in the expected cost of a trip.

Fare increases of this form impact the earnings process through three channels. First, drivers directly earn more from each trip according to the new fare schedule. Second, to the extent that rival drivers change their labor supply behavior, the level of competition and therefore the expected searching times are impacted. Third, there is a demand response: higher earnings will depress demand and lengthen expected searching times. In determining individual labor supply elasticities we only considered the impact of the first channel. Market-wide pricing changes will give rise to different work-hours elasticities because they also impact the other two channels.

To measure the elasticity of hours worked with respect to fares, we use our model and estimated parameters to simulate data under the state transitions from (1) the two months before and (2) the two months after the September 2012 fare increase. Critically, the state transitions embed all relevant information drivers need to formulate new stopping rules. Therefore, by simply observing the earnings process encountered by drivers before and after the change took effect, we can avoid modeling the search and matching process and thereby also avoid the additional restrictive assumptions that are necessary to estimate such a model (Frechette *et al.* (2019), Buchholz (2022)). By simulating driver shifts and hours worked across the *pre-* and *post*-fare hike periods, we can compute market-wide labor supply elasticities.

TABLE 8. Individual vs. Market Labor Supply Elasticity

	Baseline	I. 18% Individual Wage Increase			II. 18% Market Fare Increase		
	Hours	Hours	Elasticity	Δ Welfare	Hours	Elasticity	Δ Welfare
Overall	7.55	8.35	0.589	$65.12	7.98	0.048	$46.44
Weekday, AM	7.06	7.80	0.720	$63.36	7.73	0.079	$40.96
Weekday, PM	9.72	10.37	0.374	$40.75	9.23	-0.043	$46.77
Weekend, AM	6.25	7.22	0.862	$79.49	6.82	0.077	$47.41
Weekend PM	7.47	8.44	0.588	$90.18	8.17	0.079	$65.87

This table shows the labor hours and welfare effects of an 18% individual earnings increase (panel I) and an 18% market fare increase (panel II). Welfare is computed as $\sum_i log(exp(r_i + C(h_i|\theta_{\tau(i)}) + 1))$ for each driver-shift i and shift type $\tau(i)$.

Table 8 reports the mean work hours across all simulated shifts before and after the fare change along with implied work-hours elasticities and changes to driver welfare. We see that the overall elasticity with respect to fares is 0.048, approximately one tenth of the elasticity with respect to individual earnings. This means the benefits to a single driver are much smaller once we account for the market adjustment to the wage increase. This ratio approximately holds across different types of shifts, with the exception of Weekday PM shifts. There, drivers on average work 30 minutes *less* when the fare increases. This result captures the fact that earnings opportunities after the September 4, 2012 fare increase decline in hours worked for *only* this shift type. Moreover, at higher levels of cumulative work time, the post-fare change average spell wages are below those of the pre-fare change period. We detail these patterns in Section A.3.

We also derive welfare estimates to evaluate drivers overall benefits net of costs. Drivers are generally about 40–50% better off when faced with a personal earnings hike compared to a fare increase. Welfare increases on average by $65.12 in the former care (about 24% of the mean shift earnings). Compare this with the 17% gain in welfare following the fare hike.

The main goal of the exercise in Table 8 is to demonstrate the importance of distinguishing market-wide price and earnings variation from individual earnings variation in the assessment of labor supply elasticity. Since prices are regulated and therefore set exogenously, price changes induce both a demand and supply response as well as an equilibrium adjustment to expected search times. More generally, earnings variation may be induced by shifts in demand or supply. In all of these cases, drivers forward-looking tradeoffs and therefore labor supply decisions are impacted in unpredictable ways: demand elasticities may vary spatially, shifts in demand may be local depending on what events are taking place, etc. The mission to measure labor supply elasticity from observational data is inherently complicated by these factors. This is where our counterfactual approach can offer a clean and clear alternative; by simulating driver shifts subject to a uniform earnings increase, we are able to replicate a wage experiment without the confounding effects of equilibrium adjustments.

7. CONCLUSION

We use a comprehensive dataset of trips and work hours among New York City taxi drivers to take a new approach to a long-running question of drivers' wage elasticities by modeling taxicab

drivers' labor supply decisions as emerging from a dynamic *optimal stopping* problem. Our model explicitly assumes that drivers have standard preferences for labor and leisure, implying standard upward sloping supply curves.

We estimate our model and use it to simulate drivers' shifts. We then conduct a static reduced-form analysis of drivers' labor supply behavior that is analogous to specification used in previous literature. We demonstrate that we can replicate the same patterns in the literature, in which labor supply elasticity may appear to have a negative sign. We show that these patterns arise because of previously unexplored intra-daily dynamics in earnings per hour. In particular, drivers become less productive as they work longer, and there is also negative autocorrelation in long trips that generate apparent earnings shocks in a static framework.

Our results reconcile a twenty-five-year debate in this literature. More broadly, these findings suggest that once we account for the dynamic incentives in taxicab drivers' labor supply decisions, there is no need to add behavioral parameters to the model to explain their quitting behavior. However, we note that behavioral patterns may serve as useful heuristics for drivers that happen to coincide with the more complicated dynamic optimization problem.

Finally, our model is also capable of answering the question of what is the labor supply elasticity of New York City taxi drivers. We find individual elasticities of 0.60-0.90. These values closely match the range of averages identified in quasi-experimental studies in other settings. However, to evaluate wage policy among drivers, such as the recent fare hike among New York taxi drivers, we also have to acknowledge the equilibrium impact of wage changes as it transmits from demand elasticities as well as the spillovers from all drivers re-optimizing their search behavior. We find that the average market-level elasticity to a 18% wage hike is 0.048, or about 90% smaller than that due to a direct 18% earnings increase.

REFERENCES

ALTONJI, J. G. (1986). Intertemporal substitution in labor supply: Evidence from micro data. *Journal of Political Economy*, **94** (3, Part 2), S176–S215.

ANDERSEN, S., BRANDON, A., GNEEZY, U. and LIST, J. A. (2014). *Toward an understanding of reference-dependent labor supply: Theory and evidence from a field experiment*. Tech. rep., National Bureau of Economic Research.

ASHENFELTER, O., DORAN, K. and SCHALLER, B. (2010). A shred of credible evidence on the long-run elasticity of labour supply. *Economica*, **77** (308), 637–650.

BENKARD, C. L. (2004). A dynamic analysis of the market for wide-bodied commercial aircraft. *The Review of Economic Studies*, **71** (3), 581–611.

BESANKO, D., DORASZELSKI, U. and KRYUKOV, Y. (2014). The economics of predation: What drives pricing when there is learning-by-doing? *American Economic Review*, **104** (3), 868–897.

BRANCACCIO, G., KALOUPTSIDI, M. and PAPAGEORGIOU, T. (2020). Geography, transportation, and endogenous trade costs. *Econometrica*, **88** (2), 657–691.

BROWNING, M., DEATON, A. and IRISH, M. (1985). A profitable approach to labor supply and commodity demands over the life-cycle. *Econometrica: journal of the econometric society*, pp. 503–543.

BUCHHOLZ, N. (2022). Spatial equilibrium, search frictions, and dynamic efficiency in the taxi industry. *The Review of Economic Studies*, **89** (2), 556–591.

BURNSIDE, C., EICHENBAUM, M. and REBELO, S. (1993). Labor hoarding and the business cycle. *Journal of Political Economy*, **101** (2), 245–273.

CAMERER, C., BABCOCK, L., LOEWENSTEIN, G. and THALER, R. (1997). Labor supply of new york city cabdrivers: One day at a time. *The Quarterly Journal of Economics*, **112** (2), 407–441.

CASTILLO, J. C. (2022). Who benefits from surge pricing? *Available at SSRN 3245533*.

CHEN, M. K., ROSSI, P. E., CHEVALIER, J. A. and OEHLSEN, E. (2019). The value of flexible work: Evidence from uber drivers. *Journal of political economy*, **127** (6), 2735–2794.

CHETTY, R., GUREN, A., MANOLI, D. and WEBER, A. (2011). Are micro and macro labor supply elasticities consistent? a review of evidence on the intensive and extensive margins. *American Economic Review*, **101** (3), 471–75.

CRAWFORD, V. P. and MENG, J. (2011). New york city cab drivers' labor supply revisited: Reference-dependent preferences with rational expectations targets for hours and income. *The American Economic Review*, **101** (5), 1912–1932.

DUONG, H. L., CHU, J. and YAO, D. (2022). Taxi drivers' response to cancellations and no-shows: New evidence for reference-dependent preferences. *Management Science*.

FARBER, H. S. (2005). Is tomorrow another day? the labor supply of new york city cabdrivers. *The Journal of Political Economy*, **113** (1), 46–82.

— (2008). Reference-dependent preferences and labor supply: The case of new york city taxi drivers. *The American Economic Review*, **98** (3), 1069–1082.

— (2015). Why you can't find a taxi in the rain and other labor supply lessons from cab drivers. *The Quarterly Journal of Economics*, **130** (4), 1975–2026.

FEHR, E. and GOETTE, L. (2007). Do workers work more if wages are high? evidence from a randomized field experiment. *American Economic Review*, **97** (1), 298–317.

FRECHETTE, G. R., LIZZERI, A. and SALZ, T. (2019). Frictions in a competitive, regulated market: Evidence from taxis. *American Economic Review*, **109** (8), 2954–92.

HAGGAG, K., MCMANUS, B. and PACI, G. (2017). Learning by driving: Productivity improvements by new york city taxi drivers. *American Economic Journal: Applied Economics*, **9** (1), 70–95.

HARTWELL, A., METZ, T. and DIGIOVANNI, J. (2022). 11/15/22 commission meeting.

HECKMAN, J. and MACURDY, T. (1980). A dynamic model of female labor supply. *Review of Economic Studies*, **47** (1), 47–74.

LAGOS, R. (2000). An alternative approach to search frictions. *Journal of Political Economy*, **108** (5), 851–873.

— (2006). A model of tfp. *The Review of Economic Studies*, **73** (4), 983–1007.

MACURDY, T. E. (1981). An empirical model of labor supply in a life-cycle setting. *Journal of political Economy*, **89** (6), 1059–1085.

OETTINGER, G. S. (1999). An empirical analysis of the daily labor supply of stadium vendors. *Journal of political Economy*, **107** (2), 360–392.

PAKES, A. (1986). *Patents as options: Some estimates of the value of holding European patent stocks*. Tech. rep., National Bureau of Economic Research.

PETTERSON, M. S. (2022). Estimation of a latent reference point: Method and application to nyc taxi drivers, working Paper.

ROSAIA, N. (2023). *Competing Platforms and Transport Equilibrium: Evidence from New York City*. Tech. rep., mimeo, Harvard University.

ROTEMBERG, J. J. and SUMMERS, L. H. (1990). Inflexible prices and procyclical productivity. *The Quarterly Journal of Economics*, **105** (4), 851–874.

RUST, J. (2019). Has dynamic programming improved decision making? *Annual Review of Economics*, **11** (1), 833–858.

RYAN, S. P. (2012). The costs of environmental regulation in a concentrated industry. *Econometrica*, **80** (3), 1019–1061.

THAKRAL, N. and TÔ, L. T. (2021). Daily labor supply and adaptive reference points. *American Economic Review*, **111** (8), 2417–43.

ONLINE APPENDIX

APPENDIX A. DATA: ADDITIONAL DETAILS

A.1. **Data Cleaning and Preparation.**

A.1.1. *Data Cleaning.* We begin with raw data obtained from the New York City Taxi and Limousine commission consisting of all yellow taxi trip and fare data from 2012. The raw files consist of XXX observations. To obtain a structure suitable for estimating driver labor supply behavior, we throw out any data that appears contaminated by severe measurement error, key missing information, or highly unusual patterns.

We closely follow previous work using the same TLC dataset to prepare our data for the analysis of shifts (e.g., Haggag *et al.* (2017), Thakral and Tô (2021)). We start by cleaning the raw TLC data of obvious measurement errors or apparent extraneous data produced by duplicate or false entries or those data produced from electronic testing. The data are organized trip-by-trip, where we observe medallion identifiers and exact pickup and drop-off date-times. Our first step is to establish criteria for the change of shifts. To do this, for every driver we measure the time between trips and, for trip times with gaps beyond five hours, we define a change of shift (See discussion in Section 2). Next, we classify every driver-shift with a unique identifier. There are 1,408,646 driver-shifts in our full sample. When we identify a potentially erroneous trip or data problem, we flag the entire shift as having a problem and drop it from our analysis sample.

To begin, we flag the following problematic observations:

(i) Duplicates on medallion and pickup date-time

(ii) Missing or zero entry for trip distance or trip duration

(iii) Trip duration far less than feasible for trip distance

(iv) Trip distance far longer than feasible for trip duration

(v) Trip fare less than minimum TLC prices for normal fares (indicated as fare code=1)

(vi) Trip times less than 10 seconds

(vii) Trip times less than 60 seconds with fare greater than $10

(viii) Trips from Manhattan to JFK Airport with trip duration less than ten minutes

(ix) Trips from Manhattan to JFK Airport with trip distance less than 10 miles

(x) Trips with pickup time occurring before the previous trip's drop-off time

(xi) Latitude or longitude of trip could not be mapped to a destination within New York City or Newark Airport

36

Next, we flag the following problematic shifts:

(i) Shifts with more than one car per driver within a shift

(ii) Shifts with total duration longer than 18 hours or shorter than 2 hours

(iii) Shifts with 3 or fewer trips in total

We drop shifts with the above errors, leaving us with a data set of 8,220,299 observations, 30,231 drivers, and 444,317 unique shifts across July 1 to September 3, 2012.

A.1.2. *Predicting Medallion Types.* The TLC issues different types of medallions with different restrictions that may impact driver incentives. Although we do not have data on medallion types, we screen medallions for patterns that indicate a higher likelihood of being *fleet* medallions (subject to higher levels of minimum usage and often stringent turnover hours) vs. owner-operator medallions (which have less onerous requirements). Our screen is constructed as follows:

(i) Number of drivers per medallion less than 4

(ii) Number of trips per medallion greater than 200

The first criterion checks that the taxi medallion is only utilized by a small number of individuals, for example licensed individuals within a family. The second criterion ensures that the small number of individuals is not a consequence of scant usage of the medallion. Although this screen is simple and coarse, it predicts well the probability that the medallion will be used during the witching hour, between 4–5pm, in which fleet medallions turn over to the evening shift. The benefit of using this screen without incorporating the witching hour directly is that some owner-operators are only active in the evening shift, for which there is no equivalent of the witching hour.

A.2. **Evidence that Effort is Independent of Earnings.** In this section we investigate whether taxi drivers engage in higher (or lower) "effort" when they perceive earnings opportunities to be high (or low). We specifically characterize effort as the extent to which drivers can choose their earnings per hour across some support by paying an extra cost. Endogenous effort, or effort that varies with demand or supply shocks, would imply that quantities such as earnings per hour, wage spells, and average wages are all equilibrium objects that cannot be imposed as policy counterfactuals.

To test for endogenous effort, we evaluate whether drivers who, through lucky draws of trips, experience consistently higher or lower earnings relative to other drivers within the first four hours of the same shift (and who therefore perceive higher or lower overall demand) achieve different expected payoffs in the subsequent four hours. The question is whether drivers who perceive high demand but in reality were lucky will adjust their effort levels to be more productive later in the day. We evaluate this via a

regression that compares, among those drivers who work at least eight hours, the expected earnings of a driver in hours four through eight conditional on their first four hours. To account for the fact that there are predictable hourly patterns of earnings, we control for the hour in which the drivers' shifts begin.

Table 9 reports the results. We see that, conditional on four hour earnings, each dollar earned in the first four hours is associated with $0.031 additional earnings in the second four hours. In other words, take a one standard deviation increase in hours 0-4 income (net of date, driver and starting hour controls) of $31.87. A 1 SD increase in income over this period would predict a subsequent gain of $0.99 over the next four hours. To verify that our selection of hours 4 and 8 are not driving these results, we report alternative versions of the test in each cell and see similar results.

We interpret these results to imply that drivers' day to day success with earnings, insofar as it shapes their beliefs about future earnings within the same day, does not lead to meaningful changes in their ability to earn income compared to other drivers in the same market. Effort is difficult to measure, but these results give us some confidence that it is not driving the earnings schedules of drivers.
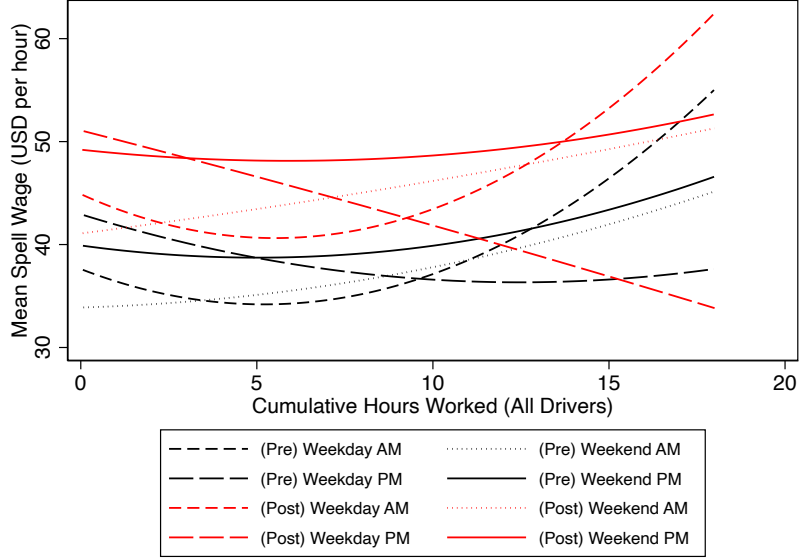
TABLE 9. Evidence for Exogenous Effort

| | Effect of 1 SD increase in earnings by: | | |
	Fourth Hour	Fifth Hour	Sixth Hour
E[Addt'l Earning] by Hour 7	$0.83	$0.22	$-0.02
E[Addt'l Earning] by Hour 8	$0.99	$0.35	$-0.01
E[Addt'l Earning] by Hour 9	$1.32	$0.62	$-0.00

This table shows the predicted change in earnings by the indicated hour in each row as a function of a one standard deviation growth in expected earnings by the hour indicated in each column. These values are obtained by regressing each row variable on shift income at the hour on each column along with date, driver and starting hour fixed effects.

A.3. **The Evolution of Average Spell Wage by Cumulative Hours and Shift Types.** Figure 3 shows how average spell wages evolve with cumulative work hours by shift type. This figure does not directly describe hourly patterns in spell wage, because shift start times are distributed across several hours in the day and evening shifts. Its purpose, however, is to reconcile Table 8 which shows that, following an 18% fare increase, work hours decreased for Weekday PM drivers. In Figure 3 we see that this category is the only one in which earnings opportunities decline relative to the pre-fare change period.

A.4. **Static Approaches to Recover Wage Elasticities.** Table 10 shows the results of elasticity regression of the form of Camerer *et al.* (1997) and further analyzed (and critiqued) in Farber (2005). Each specification regresses $log(hours)$ on $log(wage)$, where "hours" refers to the cumulative time worked by a driver upon

FIGURE 3. The Time Path of Earnings by Cumulative Hours Worked

This figure depicts, for each stated category of shifts, the expected spell wage earned by drivers with each level of cumulative hours worked. The expectation is fitted via regression on cumulative hours and cumulative hours squared.

quitting for the day, and "wage" refers to the average hourly earnings achieved through the day. In these regressions, authors derive a measure of labor supply elasticity as the coefficient on $log(wage)$. We instrument for a driver's own wage rate by using the 25th, 50th and 75th percentiles of average wages across all drivers in each day, along with day-of-week indicators. Note that these instruments shift driver wages but do not explicitly fix aggregate supply or demand. If driver wages are correlated with drivers' opportunity costs, then the coefficients become quite difficult to interpret even in a static context.

Specifications (1) and (2) in Table 10 reproduce the baseline IV regression implemented in Camerer *et al.* (1997) with and without driver fixed effects. Specifications (3)-(4) restricts the sample to fleet drivers and runs the same two regressions. Specifications (5)-(6) repeats this for owner-operator medallions. The results show that by using the richer and more granular TLC data we can attain the same apparent negative wage elasticities using an analogous empirical framework to the original result.

As Farber (2005) cautions, conventional (static) wage regressions, as in Table 10, are not easily interpretable in settings like this where intra-daily wages are variable. Because drivers do not face a fixed wage rate, decisions over how long to work can not be made on the basis of average earnings per hour unless that earnings rate is predictable and stationary, and thus a proxy for future earnings. Indeed, in

TABLE 10. Wage Regressions

	(1) All	(2) All	(3) Fleet	(4) Fleet	(5) Owner-Op	(6) Owner-Op
Log Wage	-0.720**	-0.570**	-0.669**	-0.511**	-0.694**	-0.583**
	(0.004)	(0.004)	(0.006)	(0.005)	(0.007)	(0.006)
Daily High Temp	-0.001**	-0.001**	-0.001**	-0.001**	-0.001**	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Precipitation	0.013**	0.010**	0.007**	0.006**	0.020**	0.017**
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Weekday	0.013**	0.021**	0.006**	0.008**	0.022**	0.036**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
PM Shift	0.055**	0.142**	0.057**	0.235**	0.041**	0.112**
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Driver FE	X	✓	X	✓	X	✓
N	2,065,186	2,065,044	1,136,140	1,135,116	929,046	928,608

January-August 2012 data. Data record the final cumulative hours and average wage earned as of the last trip of each driver-shift. Instruments are the 25th, 50th and 75th percentile across all driver wages each day. Standard Errors clustered at the driver-shift level. Asterisks (**) indicate significance at or above the 1% level.

Section 2.1 we show that earnings per hour is declining in cumulative hours worked, a pattern that will naturally give rise to negative correlation in the wage regressions. This fact alone suggests that wage regressions will face a selection problem: all else equal, drivers who decide to work longer will be less productive earners near the end of their shift.

A.5. **Hazard Models.** Farber (2005) and subsequent literature adopts hazard models instead of wage regressions in order to evaluate drivers' labor supply decisions. In contrast to data summarizing drivers' work time and earnings at the end of the shift, the hazard models use much more detailed trip sheet data to examine how drivers' decisions to quit relate to their total hours worked and cumulative income earned after each trip. These models allow for drivers to react differently to hours and income as they are accrued through a series of stochastic trip draws. A key finding of Farber (2005) is that drivers' quit decisions are mostly responsive to hours worked and there is only weak evidence that accrued income ever serves as a determinant of quitting. These effects are complemented by more nuanced tests for reference dependence in Farber (2008) and more detailed data in Farber (2015). In Section A.6 we broadly demonstrate this ambiguity by directly reporting quitting probabilities by cumulative work time and cumulative income.

Acknowledging that even reference-dependent drivers are unlikely to have ex-ante predictable reference points, Crawford and Meng (2011) and Thakral and Tô (2021) investigate the role of time-varying or adaptive behavioral effects. Most recently, Thakral and Tô (2021) use the TLC data to investigate whether cumulative income impacts drivers differently depending on when the income was earned within the day. They find robust evidence that income earned closer to the end of the driver's shift is associated with a higher likelihood of quitting, whereas income earned earlier in the day was not. They argue that this reflects a pattern of *adaptive reference points*, in which drivers overreact to positive earnings shocks in the short-run, but the overreaction effect of these surprises diminishes with additional work time. While the dataset used in Thakral and Tô (2021) is the same as ours, we nevertheless use a slightly different sample period. For sake of completeness we reproduce these patterns using our sample in Table 11.

TABLE 11. Hazard Model with Time-dependence

Shift Length	7hr	8hr	9hr	10hr	11hr	12hr
Income in hour 4	-0.0011	-.0026**	-.0024**	.0033**	.0003	-.0022
Income in hour 5	0.0062**	-.0003	-.0015	-.0049**	-.0002	-.0004
Income in hour 6	0.0059**	.0116**	-.0019	.003	-.0022	-.0029
Income in hour 7	-0.0652**	.0849**	.0225**	-.0031	-.0002	-.0004
Income in hour 8		-.0935**	.1284**	.0339**	.0028	-.0006
Income in hour 9			-.1443**	.1745**	.0378**	.0023
Income in hour 10				-.2024	.2036**	.0371**
Income in hour 11					-.2375**	.1871**
Income in hour 12						-.2191**
Controls	cumulative hours					
Driver FE	✓	✓	✓	✓	✓	✓
Dow x Hour	✓	✓	✓	✓	✓	✓
Dropoff Location FE	✓	✓	✓	✓	✓	✓
N	376,198	334,452	265,992	170,669	83,514	30,556

Jan-Mar 2012 data."7hr" column designates shifts with total duration between 7hr 20min and 7hr 40min. Standard Errors clustered at the driver-shift level. Coefficients depict the effect of increasing income within a given hour by $23, approx. 10% of the daily average income.

The hazard models estimate stopping probabilities as a function of both cumulative earnings and cumulative time worked. Since these two independent variables are separated, their independent variation is used to identify parameters. Therefore the broad pattern of decline in earnings per hour will not threaten identification as in the wage regressions. Nevertheless, these regressions allow us to emphasize the need for a dynamic model. Stopping decisions are inherently dynamic; the tradeoff between stopping now and stopping later is determined by the costs of additional work time and the

expected benefits of future earnings. Regressing this decision on past outcomes can only reveal an effect through the serial correlation between the current and future state variables. Since that serial correlation in earnings per time is actually negative in a short time horizon, there is a clear risk of mis-interpreting the effect of earnings shocks. This pattern, viewed through the lens of a model of forward-looking drivers, can explain the apparent recency bias in drivers' behavior.

A.6. **Descriptive Analysis of Quitting Probabilities.** In this section we show that our data broadly corroborates a finding of Farber (2005) that the hazard rates analysis demonstrates some ambiguity about the presence of income targets.

Table 12 provides a set of quitting probabilities by cumulative hours worked and cumulative income over a shift. This table reveals a broadly increasing pattern of increasing quit probabilities by both hour and income, although there are some interesting regions of quitting probabilities decreasing in income (eg. at $300-$400 and 8-9 hours of work).[15] These patterns are similar to those revealed by the hazard model estimates of Farber (2005), and we note that the limited apparent increase, albeit small, in stopping probability associated with accruing more income may again signal the possibility of negative earnings elasticities within some states.[16]

---

[15]The pattern comes with a caveat that the more extreme off-diagonal cells have relatively few observations despite the abundance of data, as those depicted in gray shading.

[16]Farber (2008) follows this work with additional evidence that reference-dependent preferences is likely limited.

TABLE 12. Choice Probabilities by Cumulative Earnings and Hours

Cum. Hours Worked	Cumulative Income Earned													
	$50	$100	$150	$200	$250	$300	$350	$400	$450	$500	$550	$600	$650	$700
1	0.000	0.001	0.000											
2	0.012	0.014	0.029	0.021										
3	0.017	0.014	0.021	0.046	0.040									
4	0.025	0.019	0.022	0.030	0.037	0.048								
5	0.041	0.028	0.027	0.036	0.047	0.084	0.062							
6	0.067	0.045	0.041	0.049	0.067	0.077	0.087							
7	0.094	0.077	0.072	0.073	0.087	0.096	0.100	0.093						
8	0.186	0.120	0.121	0.129	0.138	0.137	0.119	0.111	0.154					
9	0.232	0.172	0.182	0.200	0.220	0.216	0.185	0.152	0.143	0.160				
10	0.191	0.192	0.219	0.238	0.280	0.294	0.274	0.247	0.230	0.282	0.368			
11	0.333	0.179	0.232	0.232	0.251	0.276	0.287	0.301	0.315	0.372	0.509			
12	0.143	0.176	0.174	0.184	0.189	0.191	0.187	0.195	0.232	0.271	0.291	0.308		
13		0.750	0.200	0.184	0.163	0.177	0.163	0.165	0.178	0.185	0.197	0.200	0.333	
14			0.294	0.219	0.255	0.170	0.194	0.169	0.176	0.184	0.193	0.232	0.067	
15				0.304	0.305	0.243	0.221	0.219	0.213	0.236	0.219	0.235	0.259	
16				0.389	0.362	0.373	0.291	0.287	0.326	0.323	0.289	0.314	0.354	0.182
17					0.643	0.818	0.539	0.587	0.597	0.555	0.601	0.579	0.700	0.684

TLC Data from January through October, 2012. Each cell shows the fraction of time drivers in each category (of cumulative hours worked and income earned) quit for the day. Each category reflects values at or above the category label. For example, income category $100 is read as "$100-199.99" and hour category 1 is read as "1 hour 0 minutes - 1 hour 59 minutes". Gray entries denote cells with fewer than 100 observations. Entries with fewer than 10 observations are omitted.

A.7. **Expected Search Times by Borough.** To simplify value function computation, we model the value of making decisions in outer (i.e. non-Manhattan) boroughs as the value of decisions within Manhattan subject to an additional search time. This additional time accounts for the extra time spent among drivers who drop off passengers in the outer boroughs to get back to locations, largely in Manhattan or some sections of Brooklyn where there is a higher likelihood of finding a passenger.

To obtain the additional expected search time, we regress the waiting time of each spell on the lagged drop-off borough (i.e. the location in which the decision is made) as well as cumulative income, cumulative time worked, indicators for AM shift, weekday and owner-operator, as well as driver-shift

fixed effects. The coefficient on lagged drop-off borough (with Manhattan as the omitted category), reveals the anticipated extra search time required, above the expected search time in Manhattan, to obtain the next fare conditional on each location. Results are reported in Table 13.

TABLE 13. Search Time from Boroughs

	Bronx	Brooklyn	Newark, NJ	Queens	Staten Island
Add'l Search Time (min.)	20.81	11.70	31.82	24.72	35.15

This table shows the coefficients from a regression of taxi waiting times on the borough of the previous drop-off (with Manhattan as the omitted category) along with controls for cumulative hours, cumulative income, am/pm shift and weekday/weekend, plus driver-shift fixed effects. Newark, NJ is the site of Newark Liberty Airport, where TLC taxis are licensed to drop passengers for an additional fare. The coefficients on waiting time by borough are presented. Since the coefficients are relative to Manhattan, we can interpret these as expected additional search times associated with outer borough drop-offs.

A.8. **Simulation Details.** In Section 5.1 we use the estimated model to simulate driver shifts and then use those simulated data to reproduce key behavioral puzzles in the literature. For each simulated driver of type d, we draw a sequence of trips as (dr_t, dh_t) from the empirical distribution $F(dr, dh|r, h, x)$, draw a sequence of drivers' net shocks to the outside option ϵ_t, and for each decision point t together with parameter estimates, we compute whether drivers' value of quitting exceeds the value of continuing search. As with estimation, there are eight persistent driver types i and five categories of demand heterogeneity $k \in \{1, ..., 5\}$, defined as quintiles over average daily spell wages.

Simulation Steps. To simulate data we adopt the following procedure:

(i) Uniformly draw a demand quintile k. We assume drivers observe the day's demand type and have expectations that are consistent with the empirical state transition matrix for days of type k.

(ii) Each shift begins at the initial state $(0, 0)$, with zero cumulative income and time worked. For each of 20,000 simulated shifts, we draw from the empirical distribution of *realized* sequences of trips. This begins by selecting a random shift s from a driver-day of type i and demand quintile k. We take the first trip of s and simulate spells as they occurred in the data. Denote the simulated shift by $\hat{s}$.[17]

---

[17]Drawing from the realized distribution of trips ensures our simulated spells will follow the wage process outlined in Section 2.1.

(iii) At the end of each spell, drivers weigh the opportunity to quit and receive an immediate payoff against the option value of continuing to work longer and accrue additional earnings and time costs.

(iv) If our simulated shift $\hat{s}$ exceeds the set of trips observed in s, we append draws from another randomly selected shift s' given i and k by additionally matching the origin of the first draw in s' with the destination of our final observed trip in s.

(v) Finally, we combine each driver-day-type simulation together in proportion to their appearance rates in the data to assemble the simulated data set.

A.9. **Existence and Uniqueness of the Optimal Stopping Rule.**

Proof. We begin by writing out the Bellman equation representing driver's optimal stopping problem:

$$V_t^*(r_t, h_t, \epsilon_t) = \max \left\{ u(r_t, h_t) + \epsilon_t, E\left(V_{t+1}^*(r_{t+1}, h_{t+1}, \epsilon_{t+1}) \middle| r_t, h_t, \epsilon_t\right) \right\}$$

Suppose V_t^* is strictly monotone in ϵ_t. Then the optimal stopping rule can be described as follows. Let $y_t = 1$ denotes the "quiz" decision at time t

$$y_t = 1\left\{\epsilon_t \geq \epsilon_t^*(r_t, h_t)\right\}$$

where ϵ_t^* is obtained from

$$u(r_t, h_t) + \epsilon_t^* = E\left(V_{t+1}(r_{t+1}, h_{t+1}, \epsilon_{t+1}) \middle| r_t, h_t, \epsilon_t^*\right).$$

Given $\epsilon_t^*(\cdot)$ that defines the optimal stopping rule, we define the following equilibrium expected stopping probabilities:

$$q_{t,s}^*(r_t, h_t, \epsilon_t) \equiv P(\epsilon_{t+s} \geq \epsilon_{t+s}^*(r_{t+s}, h_{t+s}) | r_t, h_t, \epsilon_t), \forall s \geq 1$$

and

$$p_{t,s}^*(r_t, h_t) \equiv P(\epsilon_{t+s} \geq \epsilon_{t+s}^*(r_{t+s}, h_{t+s}) | r_t, h_t), \forall s \geq 0,$$

Where, by definition,

$$p_{t,s}^*(r_t, h_t) = E\left[q_{t,s}^*(r_t, h_t, \epsilon_t) \middle| r_t, h_t\right].$$

It is helpful to first consider a numerical solution: let $\epsilon_t^{(0)} = -\infty$. In this case, the driver always chooses to quit at each decision period. Then $q_{t,s}^{(1)}(\cdot) = 1$ for all t and $s \geq 1$, and

$$V_t^{(0)}(r_t, h_t, \epsilon_t) = \max\left\{u(r_t, h_t) + \epsilon_t, 0\right\} \geq 0.$$

Now update the optimal decision and value function as follows:

$$\epsilon_t^{(1)}(r_t, h_t) = -u(r_t, h_t)$$

and

$$V_t^{(1)}(r_t, h_t, \epsilon_t) = \max\left\{ u(r_t, h_t) + \epsilon_t, E\left(V_{t+1}^{(0)}(r_{t+1}, h_{t+1}, \epsilon_{t+1})\big|r_t, h_t, \epsilon_t\right)\right\}$$

Clearly, $\epsilon_t^{(1)} \geq \epsilon_t^{(0)}$ and $V_t^{(1)} \geq V_t^{(0)}$. Moreover, $\epsilon_t^{(2)}$ is obtained by solving $\bar{\epsilon}_t$ from the following equation

$$u(r_t, h_t) + \bar{\epsilon}_t = E\left(V_{t+1}^{(1)}(r_{t+1}, h_{t+1}, \epsilon_{t+1})\big|r_t, h_t, \bar{\epsilon}_t\right)$$

and then

$$V_t^{(2)}(r_t, h_t, \epsilon_t) = \max\left\{ u(r_t, h_t) + \epsilon_t, E\left(V_{t+1}^{(1)}(r_{t+1}, h_{t+1}, \epsilon_{t+1})\big|r_t, h_t, \epsilon_t\right)\right\}.$$

We now show $\epsilon_t^{(s+1)}(\cdot) \geq \epsilon_t^{(s)}(\cdot)$ and $V_t^{(s+1)}(\cdot) \geq V_t^{(s)}(\cdot)$ by induction: Because

$$V_t^{(s)}(\cdot) \geq V_t^{(s-1)}(\cdot)$$

It is straightforward that

$$V_t^{(s+1)}(\cdot) \geq V_t^{(s)}(\cdot)$$

Thus,

$$E\left(V_{t+1}^{(s+1)}(r_{t+1}, h_{t+1}, \epsilon_{t+1})\big|r_t, h_t, \epsilon_t^{(s+1)}\right) \geq E\left(V_{t+1}^{(s)}(r_{t+1}, h_{t+1}, \epsilon_{t+1})\big|r_t, h_t, \epsilon_t^{(s+1)}\right)$$

holds almost surely.

It follows that $\epsilon_t^{(s+1)}(\cdot) \geq \epsilon_t^{(s)}(\cdot)$. This implies the existence of an optimal stopping rule. Next we show uniqueness by contradiction. We assume the following: (i) the distribution of ϵ has full support, (ii) $u(\cdot)$ is bounded, (iii) conditions to ensure V_t^∞ to be bounded, which holds under our specification of the utility function and (iv) $u(\cdot)$ is decreasing and convex in h. Let $\tilde{V}_t$ be another equilibrium. By definition, $V_t^{(s)}(\cdot) \leq \tilde{V}_t(\cdot)$ and $\epsilon_t^{(s)}(\cdot) \leq \tilde{\epsilon}_t(\cdot)$ for all s. Therefore, $V_t^\infty(\cdot) \leq \tilde{V}_t(\cdot)$.

Let $\delta \equiv \sup_{r,h,e} \tilde{V}_t(r,h,e) - V_t^\infty(r,h,e)$. Then for an arbrtrary (r_t, h_t, ϵ_t):

$$\tilde{V}_t(r_t, h_t, \epsilon_t) - V_t^\infty(r_t, h_t, \epsilon_t)$$

$$= \max\left\{ u(r_t, h_t) + \epsilon_t, E\left(\tilde{V}_{t+1}(r_{t+1}, h_{t+1}, \epsilon_{t+1})\big| r_t, h_t, \epsilon_t\right)\right\}$$

$$- \max\left\{ u(r_t, h_t) + \epsilon_t, E\left(V_{t+1}^\infty(r_{t+1}, h_{t+1}, \epsilon_{t+1})\big| r_t, h_t, \epsilon_t\right)\right\}$$

$$\leq \max\left\{ u(r_t, h_t) + \epsilon_t, \delta + E\left(V_{t+1}^\infty(r_{t+1}, h_{t+1}, \epsilon_{t+1})\big| r_t, h_t, \epsilon_t\right)\right\}$$

$$- \max\left\{ u(r_t, h_t) + \epsilon_t, E\left(V_{t+1}^\infty(r_{t+1}, h_{t+1}, \epsilon_{t+1})\big| r_t, h_t, \epsilon_t\right)\right\} < \delta.$$

This is a contradiction.

$\square$