Kamble, Vikrant V.; Motghare, Swapnil

Working Paper

Intertemporal Elasticity of Labor Supply: Evidence from New York City Taxicab Drivers Using a New Instrument

GLO Discussion Paper, No. 1809

Provided in Cooperation with:
Global Labor Organization (GLO)

Intertemporal Elasticity of Labor Supply: Evidence from New York City Taxicab Drivers Using a New Instrument*

Vikrant V. Kamble[†]
Department of Economics, University of Delaware;
Global Labor Organization (GLO)

Swapnil Motghare
California Policy Lab, University of California Los Angeles

August 27, 2026

Abstract

We estimate the intertemporal elasticity of labor supply for New York City taxicab drivers using a new instrument: the type of taximeter installed in the vehicle. The two meter systems in use display different default tip percentages, generating plausibly exogenous variation in tip income and hourly pay across shifts. Assignment to the "high-default" meter raises hourly wages by 0.5 percent, entirely through tips, and increases shift hours by 0.9 percent, implying an elasticity of 1.7. These findings align with the standard neoclassical prediction that workers supply more hours when temporary pay rises and shed light on labor-supply behavior in flexible, schedule-setting work environments more broadly.

Keywords: labor supply; Frisch elasticity; intertemporal substitution; taxi drivers; tipping; instrumental variables

JEL codes: J22, J31, D15, D91, C26

*We are grateful to Kwabena B. Donkor for his helpful feedback. We sincerely thank Desmond Toohey, Barton Willage, Olga Gorbachev, and Maria Rosales-Rueda for their valuable comments and suggestions. We also thank Parag Mahajan for constructive feedback on an earlier draft of this paper. We are also grateful to Nilesh N. Shinde, Vikas Gawai, Prashikdivya Gajbhiye, and Sachin Badole for their suggestions. Finally, we thank the participants at the International Conference on Empirical Economics at Penn State Altoona for their insightful questions and valuable feedback during the presentation. The authors are solely responsible for any remaining errors or omissions. The authors declare that they have no conflicts of interest.

[†]Corresponding author: vvkamble@udel.edu

1 Introduction

How workers respond to changes in pay is central to labor policy and to firm behavior alike, from overtime regulation and minimum wage design to pricing by employers who rely on flexible labor. It governs the employment impact and efficiency loss of progressive income taxation (Stafford, 2018). And the question has grown more pressing as a rising share of workers set their own schedules through rideshare driving, delivery services, and freelance platforms.

The standard neoclassical model predicts that a temporary increase in pay leads workers to supply more hours. Given an hourly wage, a worker chooses how much leisure to consume, and a higher wage pulls in two directions: it makes him richer and so raises his demand for leisure, the income effect, but it also raises the opportunity cost of leisure and so encourages work, the substitution effect. A transitory increase barely moves lifetime income, so the income effect is small, the substitution effect dominates, and workers "make hay while the sun shines." The strength of that response is the Frisch elasticity of labor supply, the percentage change in hours worked for a one percent change in the wage, holding the marginal utility of wealth constant. It is the parameter that governs labor supply under transitory wage variation, it is the parameter we estimate, and Section 3 derives it formally.

Testing this prediction requires workers who genuinely decide their own hours, and most of the labor force does not. Hours are typically fixed by contract, bounded by negotiated or statutory standard workweeks, or adjusted at the employer's initiative rather than the worker's (Hunt, 1999, 2012). Where discretion is real it tends to belong to salaried and highly paid workers and to track the incentives written into their contracts (Kuhn and Lozano, 2008; Stoddard and Kuhn, 2004). The returns to supplying long hours are large enough to shape occupational choice and the skilled gender wage gap (Cortés and Pan, 2019; Cortés and Tessada, 2011; Cortes and Pan, 2013); such discretion is delegated precisely where firms cannot monitor inputs (Prendergast, 2002; Barlevy and Neal, 2019), and workers value it alongside pay itself (Lamadon et al., 2022, 2024). New York City taxi driving sits at the far end of this spectrum: no supervisor, no scheduled shift length, no penalty for stopping early. The margin the theory concerns is one these drivers can actually adjust.

Early research in settings of this kind produced conflicting results. Work on taxi drivers in New York City (Camerer et al., 1997) and Singapore (Chou, 2002) documented negative labor supply responses consistent with income targeting: drivers quit early on high earning days and extended hours on low earning ones. Later work using richer data and stronger identification finds the opposite, though always for specific groups of workers. Studies of baseball stadium vendors (Oettinger, 1999), bicycle messengers in Zurich (Fehr and Goette, 2007), New York City taxi drivers (Farber, 2015), Florida lobster fishermen (Stafford, 2015), and Uber drivers in Boston (Angrist et al., 2021) consistently show that hours rise when temporary earnings opportunities improve. On balance the recent literature supports the neoclassical prediction of positive short run elasticities, but the disagreement over both sign and magnitude has not been settled by a design that isolates the wage from everything moving alongside it.

In this paper we revisit whether workers supply more hours when wages are higher, using data on New York City yellow cab drivers and a previously unused source of exogenous wage variation: the taximeter installed in each vehicle. The two meter systems present different default tip options, and because default prompts shape tipping, meter type generates predictable differences in tip income and so in hourly pay. Crucially, most taxis are rented and drivers take whichever vehicle is available, so the same driver often works different meter types across shifts. That institutional feature produces plausibly exogenous within driver variation in shift level pay, which we use to compare how a driver adjusts hours from one shift to the next. Because the typical shift level wage change is small it leaves lifetime wealth essentially unchanged, and driver fixed effects absorb the driver specific marginal utility of wealth, following MaCurdy (1981).

We estimate that assignment to the meter displaying higher default tips raises hourly wages by roughly 0.5 percent, entirely through tips, and hours worked by 0.9 percent, implying a short run labor supply elasticity of 1.7. This exceeds most values in the literature, a difference that may reflect the scheduling flexibility taxicab drivers enjoy. The estimates are robust to alternative definitions of a work shift, though somewhat sensitive to the inclusion of additional controls. Weak first stage variation limits several of our heterogeneity analyses, leaving the elasticity unidentified for owners, for drivers who work mainly by day, and for the most experienced drivers. Where identification is strong, renters display elasticities close to the overall estimate and drivers who work mainly at night display elasticities near zero; in no subsample do we find evidence of a negative short run elasticity.

This paper makes three contributions. First, we introduce a new source of exogenous variation in the wage. Estimating labor supply elasticities from observed wages is difficult because demand shocks raise wages and hours simultaneously, because long shifts may themselves depress hourly earnings through fatigue, and because hours appear in the denominator of the constructed wage, producing division bias toward negative estimates (Borjas, 1980). Taximeter assignment is unrelated to all three, and the consequence is visible in our estimates: the OLS elasticity is -0.267, while the same specification instrumented by meter type yields $+1.7$. Second, the institutional structure of our setting supplies a partial test of the exclusion restriction rather than only an argument for it. Because the two meter systems compute fares identically and differ solely in the default tip options they display, the instrument should shift tip earnings and leave fare earnings untouched. It does: the effect on hourly fare revenue is a precisely estimated zero, while hourly tip revenue rises by 7.4 percent. An instrument that admits this kind of check is uncommon.

Third, our results show that a default operating on consumers propagates into the labor supply of workers. Prior work establishes that default tip prompts change what passengers pay (Haggag and Paci, 2014); we show that this behavioral response travels downstream and changes how long drivers choose to work. The mechanism matters beyond taxis, since rideshare and delivery platforms routinely set default tip percentages and thereby influence worker earnings without adjusting base rates. Taken together, our finding that drivers extend shifts when transitory pay rises aligns with the neoclassical model and with recent empirical evidence. We do not read our results as evidence that reference dependent preferences or income targeting never arise; they show instead that in this setting behavior is well described by neoclassical incentives, and that where workers hold real discretion over their schedules, they act on short term financial incentives in the way standard theory predicts.

Our estimate also belongs to a literature in which reported magnitudes vary widely, and much of that variation reflects a basic point: what a design recovers depends on the margin, the horizon, and the nature of the wage variation exploited. Evidence from tax credits, transfer programs, retirement incentives, and statutory wage floors finds responses concentrated on the decision to work rather than on hours, with magnitudes that shift with the population and the policy (Eissa and Hoynes, 2006, 2004; Kuhn and Riddell, 2010; Hoynes, 2000; French and Jones, 2012; Gai et al., 2025; Giuliano, 2013; Bailey et al., 2021). None of these identifies the Frisch elasticity, because the wage changes are permanent or fully anticipated and so carry income effects. The closest analogue to our parameter is Martínez et al. (2021), who exploit two year income tax holidays staggered across Swiss cantons and recover an intertemporal elasticity of 0.025 overall, rising to 0.25 among those who work for themselves, with no extensive margin response at all. Their variation is anticipated, spans two years, and reaches workers whose weekly hours are largely set by employers; ours is realized within a single shift among workers who can stop driving at any moment. Read together, the two estimates suggest that the intertemporal elasticity is less a single structural constant than a function of how quickly hours can be adjusted. Measurement matters too: conventional hours series are unreliable enough to overturn long run conclusions about productivity growth (Siegenthaler, 2015), whereas our trip level data time each shift to the second, limiting the division bias identified by Borjas (1980).

Two limits on scope deserve emphasis. Our estimate applies to the intensive margin, hours worked conditional on a driver beginning a shift, rather than the extensive margin of whether he works at all. This follows from the industry and from our data: we observe a driver only once a shift has begun, and renters, who make up roughly 91 percent of our sample, pay a fixed rental fee upfront before learning which taximeter they have been assigned, so participation is largely settled before the wage variation we exploit is realized. And because our sample consists of taxicab drivers with substantial flexibility over hours, the magnitude of the estimated elasticity should not be generalized to the broader workforce. Neither limit blunts the relevance of the results for flexible work arrangements, overtime policy design, and app based labor markets, where discretion over hours is precisely the feature at issue.

The remainder of the paper is organized as follows. Section 2 provides the institutional background and Section 3 sets out the conceptual framework. Section 4 describes the data, sample construction, and key variables, and Section 5 presents the empirical strategy. Section 6 reports the main results, Section 7 a series of robustness checks, and Section 8 treatment effect heterogeneity. Section 9 discusses the findings and their broader implications, and Section 10 concludes.

2 Institutional details

Taxis in New York City operate under the regulations of the Taxi and Limousine Commission (TLC), which controls both the number of taxis through the medallion system and the licensing of drivers. The TLC also sets fare rules, including the base fare, surcharges, and allowable extras. Two types of taxis serve the city, yellow cabs and green cabs, and our analysis focuses on yellow cabs in 2010. At that time, roughly 12,000 medallions were in circulation, about 40 percent owned by individuals and the remainder by companies. Company-owned medallions are typically rented out one shift at a time, usually for 12 hours, while individual owners may both drive their own vehicles and rent them out for additional shifts.

Each taxi is equipped with a taximeter, the device that displays the starting fare, calculates the total charge, and processes credit-card payments. Two systems were in use during our study period: VeriFone Transportation Systems (VTS) and Creative Mobile Technologies (CMT). Owners, either individuals or companies, choose which system to install, but both calculate fares identically, and the TLC inspects meters every three to four months to ensure accuracy. Drivers cannot alter meter functionality; installation and maintenance are handled by TLC-authorized shops, and tampering can result in license revocation and fines.

Drivers who rent taxis do not select the vehicle they use. At fleet garages, drivers queue and are assigned whichever taxi becomes available, meaning a driver may use a VTS-equipped taxi one day and a CMT-equipped taxi the next. Rental fees are fixed, capped at $105 for a standard 12-hour day shift during the study period, and drivers face no minimum-hours requirement or penalty for ending a shift early. Although drivers may choose not to work after renting a taxi, most do so because the rental fee is paid upfront.

3 Conceptual Framework

This section derives the estimating equation of Section 5 from a life cycle model of labor supply and shows that the coefficient of interest is the Frisch elasticity. It then nests the leading alternative account, under which the same wage variation carries a sharp and opposite prediction, so that the sign of our estimate becomes a test between the two.

3.1 Setup

A driver d works a sequence of shifts indexed by s and chooses consumption c_{ds} and hours h_{ds} to solve

$$\max\ \mathbb{E}_0 \sum_{s=0}^{S} \beta^s \left[u(c_{ds}) - \varphi_{ds} v(h_{ds})\right] \quad \text{s.t.} \quad A_{d,s+1} = (1+r)\left(A_{ds} + w_{ds}h_{ds} - R_{ds} - c_{ds}\right), \quad A_{d,S+1} = 0, \tag{1}$$

where w_{ds} is the realized hourly wage, R_{ds} the lease payment, and φ_{ds} a shift specific shifter of the disutility of driving that absorbs weather, congestion, fatigue, and anything else making an hour behind the wheel more or less costly. We assume $u' > 0$, $u'' < 0$, $v' > 0$, and $v'' > 0$. Letting λ_{ds} denote the marginal utility of wealth, the first order condition for hours is

$$\varphi_{ds} v'(h_{ds}) = \lambda_{ds} w_{ds}. \tag{2}$$

The driver equates the marginal disutility of an additional hour to its marginal value, the wage scaled by the marginal utility of wealth.[1]

3.2 The Frisch elasticity and the estimating equation

Take

$$v(h) = \frac{h^{1+1/\gamma}}{1+1/\gamma}, \qquad \gamma > 0, \tag{3}$$

so that $v'(h) = h^{1/\gamma}$. Substituting into Equation 2 and taking logarithms gives

$$\ln h_{ds} = \gamma \ln w_{ds} + \gamma \ln \lambda_{ds} - \gamma \ln \varphi_{ds}, \qquad \gamma = \left.\frac{\partial \ln h_{ds}}{\partial \ln w_{ds}}\right|_{\lambda_{ds},\varphi_{ds}}, \tag{4}$$

so that γ is the Frisch elasticity, the response of hours to a wage change holding the marginal utility of wealth fixed. This is the relevant parameter here because the variation we exploit is transitory and small. A driver assigned the high default meter earns roughly a dollar and a half more across an entire shift, an amount that changes lifetime wealth negligibly. We therefore assume that the

[1] Throughout, λ_{ds} denotes the multiplier on the shift s budget constraint rescaled into current shift utility units, so that the discount factor does not appear explicitly in Equation 2. Because the lease payment R_{ds} is a lump sum that does not vary with hours, it drops out of this condition and affects the hours decision only through the level of λ_{ds}. We impose no borrowing constraint; a binding constraint would raise λ_{ds} for the drivers it affects without altering the sign of the wage response in Equation 4.

induced change in λ_{ds} is negligible, so that the substitution effect accounts for the relevant short run response. The same assumption delivers the economic content of our exclusion restriction: meter assignment can shift hours through w_{ds} in Equation 2, but it is too small to materially affect the marginal utility of wealth and has no direct reason to enter φ_{ds}.

The Euler equation implies

$$\lambda_{ds} = \beta(1+r)\,\mathbb{E}_s\left[\lambda_{d,s+1}\right], \tag{5}$$

so that, iterating forward,

$$\ln \lambda_{ds} = \ln \lambda_{d0} - s\ln[\beta(1+r)] + \xi_{ds}, \tag{6}$$

where λ_{d0} denotes the driver's marginal utility of wealth at the start of the sample and collects the component determined by initial wealth, preferences, and expectations (MaCurdy, 1981), and ξ_{ds} accumulates the forecast errors realized up to shift s and vanishes under perfect foresight. The driver fixed effect α_d absorbs $\gamma \ln \lambda_{d0}$.[2]

We write the disutility shifter as

$$\ln \varphi_{ds} = -\gamma^{-1}\left(\delta_s + \theta_s + X_{ds}\beta'\right) + \widetilde{\varepsilon}_{ds}, \tag{7}$$

where δ_s and θ_s denote fixed effects for the calendar date and the start hour of shift s, so that date effects capture citywide conditions, hour of day effects capture systematic variation in the cost of driving over the day, and X_{ds} captures payment composition, trip geography, and other observable shift characteristics. Substitution into Equation 4, absorbing $\gamma \ln \lambda_{d0}$ into the driver fixed effect and setting aside the negligible shift order drift, yields

$$\ln h_{ds} = \beta_1 \ln w_{ds} + X_{ds}\beta' + \alpha_d + \delta_s + \theta_s + \epsilon_{ds}, \qquad \beta_1 = \gamma, \tag{8}$$

which is the specification we take to the data in Section 5, with h_{ds} and w_{ds} corresponding to hours worked and the hourly wage of driver d on shift s. Thus, under the maintained assumption that

[2] Two terms remain. The first is a driver specific linear drift in shift order; because s counts a driver's own shifts rather than calendar time, date effects do not sweep it out, but with $\beta(1+r)$ close to unity and a sample spanning a single year the implied drift is negligible relative to shift to shift variation in wages. The second is $\gamma\xi_{ds}$, which enters the residual. Since revisions to expected lifetime wealth may correlate with realized earnings, this is a further reason the wage in Equation 8 cannot be treated as exogenous, and it is addressed by the instrument rather than by the fixed effects.

the meter induced wage change has a negligible effect on the marginal utility of wealth, β_1 has the interpretation of a short run Frisch rather than Marshallian elasticity.

3.3 Reference dependent preferences

The neoclassical account is not the only candidate, and the framework above nests its main rival. Suppose the driver evaluates shift earnings $y_{ds} = w_{ds}h_{ds}$ against a reference level T_d, so that period utility becomes

$$u(c_{ds}) - \varphi_{ds}v(h_{ds}) + \mu(y_{ds} - T_d), \tag{9}$$

where μ is increasing and kinked at zero: an additional dollar of shift earnings carries marginal utility $\eta > 1$ while the driver remains short of the target and marginal utility one once he has passed it. This asymmetry encodes loss aversion around the reference point, and the neoclassical case is recovered at $\eta = 1$, where the kink vanishes. The daily income targeting hypothesis originates with Camerer et al. (1997); the structural formulation we adopt here follows Farber (2008).

For a driver far from the target, the marginal utility of earnings is locally constant and the model retains a positive wage hours relationship. The two accounts part company most sharply for drivers whose hours are constrained by their earnings targets. The kink gives such a driver a strong reason to keep working until earnings reach T_d and a markedly weaker one thereafter, and a higher wage brings him to that point sooner. In the limiting case in which he stops exactly at the target, $w_{ds}h_{ds} = T_d$, so that

$$h_{ds} = \frac{T_d}{w_{ds}}, \tag{10}$$

and therefore

$$\frac{\partial \ln h_{ds}}{\partial \ln w_{ds}} = -1. \tag{11}$$

More generally, the aggregate response under reference dependent preferences depends on the share of drivers whose hours are constrained by earnings targets and on the wage response of drivers away from the target. The aggregate elasticity can therefore be negative when the targeting response is sufficiently prevalent, and it approaches -1 in the limiting case in which all drivers stop at their earnings targets. The neoclassical model, in contrast, predicts a positive Frisch elasticity under the maintained assumptions. The sign of the estimated wage response therefore provides a useful test between the two mechanisms under the wage variation we exploit.

We note that Farber (2008), having estimated this model on data for New York City taxi drivers, concludes against it: estimated reference levels vary considerably from one day to the next, and most shifts end before the estimated reference level is reached. Our design offers an independent route to the same question. Because our instrument moves the wage without, under the exclusion restriction, directly moving the disutility of driving or the driver's underlying wealth, the test isolates wage variation from contemporaneous changes in demand conditions, and the sign of the response does not require the reference level to be recovered at all. A positive estimate is consistent with the neoclassical labor supply mechanism, whereas a negative estimate is consistent with a sufficiently strong earnings targeting mechanism.

4 Data

We use data on all trips taken by New York City yellow cabs in 2010, a subset of the dataset used in Motghare (2021).[3] These data are generated by the taximeters installed in each taxi, which record detailed information on every trip. For each trip, we observe an encrypted driver license number and a medallion ID, allowing us to identify the driver and the taxi used. The data also report trip start and end times, passenger counts, fare amounts, payment method, tip amounts, and the taximeter type installed in the vehicle.[4] In 2010, the dataset covers approximately 36,000 drivers operating 12,534 medallions, of whom 90.6 percent are renters. The distribution of taximeter types is roughly even: 47.6 percent of taxis use CMT and 52.4 percent use VTS.

We aggregate trips into work shifts, defined as the period during which a driver takes a taxi out to work. Following Farber (2015), we group trips into shifts based on gaps between consecutive trips. Shift-level hourly wages are computed as total shift earnings divided by total shift duration. Drivers are classified as renters or owners using the method in Motghare (2021). The final dataset contains roughly six million shifts completed by about 35,000 drivers. More details on the shift-level aggregation can be found in Appendix A.

[3]We focus on data from 2010 since 2011 and 2012 show a substantially reduced default tip differential between vendors, limiting our identifying variation. We exclude September 2010 because a non-random subset of drivers is missing from the data in that month.

[4]We do not observe cash tips. The taximeter only records tips when passengers pay by credit card and cash tips cannot be recorded. A small number of cash tips that are observed in the data are likely due to measurement error from treating card tips as cash. More details are mentioned in Appendix A.

Table 1 reports summary statistics for shifts by taximeter type. About 93 percent of CMT shifts and 87 percent of VTS shifts are driven by renters. Average shift duration is similar across systems: 9.41 hours for CMT and 9.38 hours for VTS. Hourly wages and tip wages are also comparable. The geographic distribution of trips is nearly identical across meter types: both groups work predominantly in Manhattan, with smaller shares in the Bronx, Brooklyn, and Queens. Overall, the similarity of shifts across CMT and VTS supports the view that taximeter type is determined by the fleet or company a driver works for, rather than by driver characteristics or choices.

5 Empirical Strategy

We measure how taxi drivers respond to changes in their hourly wage. That response is commonly summarized by the labor supply elasticity, defined as the percentage change in hours worked associated with a one-percent change in wages. To estimate this elasticity, we use the following regression model.

Our baseline specification relates hours worked to hourly wages at the driver–shift level:

$$\ln(\text{hours}_{ds}) = \beta_1 \ln(\text{wage}_{ds}) + X_{ds}\beta' + \alpha_d + \delta_s + \theta_s + \epsilon_{ds} \tag{12}$$

where $\ln(\text{hours}_{ds})$ is the natural logarithm of hours worked by driver d in shift s, and $\ln(\text{wage}_{ds})$ is the natural logarithm of the hourly wage earned by driver d on that shift. The coefficient of interest, β_1, is the labor-supply elasticity: a positive value indicates that drivers work more when wages are higher, while a negative value indicates the opposite.

The vector X_{ds} includes the share of trips paid by credit card and the share of trips originating in each borough (Manhattan, Bronx, Brooklyn, and Queens, with Staten Island omitted). Payment composition matters because credit-card trips carry systematically different tipping behavior, and borough shares absorb the fact that a shift spent in Manhattan faces a different fare-per-hour environment than one spent in Queens.

The three sets of fixed effects address distinct sources of omitted-variable bias. Driver fixed effects (α_d) absorb everything permanent about a driver: skill at locating passengers, tolerance for long shifts, target earnings, and any fixed financial obligation such as a lease payment. Absent them, the regression would compare fast drivers who both earn more per hour and are willing to stay out longer against slow drivers who do neither, generating a spurious positive elasticity that reflects driver heterogeneity rather than any behavioral response to wages. Including α_d restricts identification to variation within a driver across shifts, so the question becomes whether a given driver works longer on his own higher-wage shifts.

Date fixed effects (δ_s) absorb citywide shocks common to all drivers on a given day: rain, snow, a subway outage, a holiday, a parade, or a fare change. These shocks move wages and hours simultaneously and in ambiguous directions, so leaving them in the error term contaminates β_1 with demand conditions rather than labor supply. Rain, for instance, raises the wage by tightening the market while also raising the cost of driving, so the observed hours response confounds the wage effect with the disutility of working in bad weather.

Hour-of-day fixed effects (θ_s) absorb the fixed daily rhythm of demand. Wages are predictably high during the morning and evening rush and predictably low in mid-afternoon, while shift length is jointly determined by the same clock: shifts that begin in the early morning differ in duration from those that begin in the evening for reasons that have nothing to do with the wage. Without θ_s, β_1 would partly recover this mechanical correlation between the time of day a shift starts and how long it runs. Conditioning on the starting hour means the comparison is between shifts that begin at the same point in the daily cycle but happen to draw different wages. Standard errors are clustered at the driver level, since repeated shifts by the same driver are unlikely to have independent errors.

Even after controlling for observable factors and driver fixed effects, hourly wages are likely correlated with the error term ϵ_{ds}, creating several sources of endogeneity. First, wages may be correlated with unobserved determinants of hours worked, generating omitted-variable bias. For example, when demand is high, drivers earn more per hour, largely through higher tips, and also choose to work longer because business conditions are favorable. Without fully accounting for demand, the resulting correlation between wages and hours would lead us to overstate the true responsiveness of labor supply. Second, hours worked may influence wages, producing simultaneity bias. Drivers who work very long shifts may become fatigued and provide lower-quality service, reducing tips and hourly earnings. Similarly, drivers working late at night may encounter different tipping patterns. This reverse causality makes it difficult to isolate the causal effect of wages on hours. Third, measurement error in hourly wages introduces division bias. Because hourly wage is computed as total revenue divided by hours worked, any measurement error in hours appears in both the dependent variable and the wage measure. This mechanical correlation (Borjas, 1980) biases the estimated elasticity toward negative values, even when the true elasticity is positive. For these reasons, Ordinary Least Squares (OLS) estimates of β_1 are biased, motivating the use of an instrumental-variables strategy.

To address these sources of endogeneity, we use an instrumental-variables (IV) strategy that exploits the type of taximeter installed in the taxi as an instrument for hourly wages. Taximeter type affects the wage earned during a shift through two mechanisms: (i) the default tip options displayed to passengers and (ii) the fare base used to compute those default options. During our study period, both taximeter systems calculated fares identically, but they differed in how they presented tip suggestions. CMT meters displayed default tips of 15, 20, and 25 percent of the base fare only, whereas VTS meters displayed 20, 25, and 30 percent of the base fare plus taxes and surcharges. As a result, for the same trip, passengers using a VTS-equipped taxi were shown systematically higher default tip amounts. Prior research (Haggag and Paci, 2014) documents that default tip prompts meaningfully influence actual tipping behavior. Consequently, identical trips completed with different taximeters generate different tip earnings, producing exogenous variation in hourly wages. This variation forms the basis of our instrumental-variables approach.

We estimate an instrumental-variables model that uses taximeter type as an instrument for hourly wages. The first stage captures how taximeter type affects wages:

$$\ln(\text{wage}_{ds}) = \pi_1 \text{VTS}_{ds} + X_{ds}\pi' + \alpha_d + \delta_s + \theta_s + \nu_{ds} \tag{13}$$

where VTS_{ds} equals 1 if the taxi is equipped with a VTS meter and 0 if it is equipped with a CMT meter. The coefficient π_1 measures the wage difference associated with VTS relative to CMT. We expect $\pi_1 > 0$ because VTS meters display higher default tip percentages. In the second stage, we use the predicted wage from Equation 13 to estimate the causal effect of wages on hours worked:

$$\ln(\text{hours}_{ds}) = \beta_1 \widehat{\ln(\text{wage}_{ds})} + X_{ds}\beta' + \alpha_d + \delta_s + \theta_s + \varepsilon_{ds} \tag{14}$$

where $\widehat{\ln(\text{wage}_{ds})}$ is the predicted wage from Equation 13. The coefficient β_1 from this second stage gives us the causal effect of wages on hours worked. This estimate uses only the variation in wages caused by taximeter type, not the endogenous variation from demand, fatigue, or measurement error.

For the instrumental-variables strategy to recover the causal effect of wages on hours worked, two conditions must hold. The first is instrument relevance: taximeter type must affect hourly wages. We assess this using the first-stage regression in Equation 13, where a statistically significant and sufficiently large π_1 (with an F-statistic above 10) indicates a strong instrument. We further decompose hourly wage into fare wage (fare revenue per hour) and tip wage (tip revenue per hour) and show that taximeter type is correlated only with tip wage, not fare wage. This provides additional validation, since the two taximeters calculate fares identically but differ in the default tip prompts they display.

The second condition is the exclusion restriction: taximeter type must affect hours worked only through its effect on wages. In other words, VTS meters should influence labor supply solely because they induce higher tip earnings. A violation would occur if VTS-equipped taxis differed from CMT-equipped taxis in other ways that affect hours, for example, if VTS taxis were newer or more comfortable, encouraging longer shifts. Although we do not observe vehicle characteristics directly, institutional features of the market suggest that taxis differ only in their meter vendor, and the comparison of observables in Table 1 supports this claim.

Another potential violation arises if taximeter type correlates with driver fixed effects, which could occur if drivers systematically choose higher-paying meters. Several considerations argue against this. First, the average difference in hourly tip wage between VTS and CMT is modest—about \$0.20 per hour, or \$1.60 over an eight-hour shift—unlikely to generate strong incentives for selection. Second, renter-drivers are assigned taxis on a first-come, first-served basis among vehicles returning from previous shifts, severely limiting their ability to choose. Third, even if drivers preferred VTS, the limited supply of taxis would constrain their ability to act on that preference.

We also examine the data directly for evidence of selection (Appendix D). By tracking individual drivers' sequences of meter types, we study whether drivers disproportionately choose VTS after previously driving CMT. We find no such pattern. Market shares remain stable at roughly 50–50 throughout the year, indicating that vendor assignment is determined by institutional factors rather than driver optimization. This stability supports the exclusion restriction and the validity of our identification strategy.

6 Results

We begin by establishing instrument relevance. Figure 1 shows that, across all trips, the average ratio of tip amount to fare amount is higher for VTS taximeters than for CMT taximeters. Because tips constitute roughly 20–25 percent of total fare revenue, this difference indicates that taximeter type is systematically related to tip earnings. This pattern suggests that the instrument is likely correlated with hourly wages, which we examine directly in the first-stage analysis that follows.

Table 2 examines how taximeter type affects different components of driver earnings while controlling for driver, date, and hour-of-day fixed effects, as well as the proportion of credit-card trips and pickup locations. Model 1 reports the first stage of the IV specification, with hourly wage as the dependent variable. VTS shifts earn approximately 0.5 percent more than CMT shifts, and the corresponding F-statistic of 31.99 exceeds the conventional threshold of 10, confirming that taximeter type is a strong instrument. Model 2 uses hourly fare-only wage (fare revenue per hour) as the dependent variable. The coefficient on VTS is precisely estimated at zero, consistent with the fact that both taximeters compute fares identically. Model 3 uses hourly tip-only wage (tip revenue per hour) as the dependent variable and shows that VTS shifts receive 7.4 percent higher credit-card tips. Because default tip prompts affect credit-card tipping most strongly, this pattern aligns with our identification argument. Taken together, these results demonstrate that taximeter type is correlated with hourly wage through its effect on tip earnings but not fare earnings, allowing us to transparently isolate the source of exogenous wage variation. Model 4 uses log hours as the dependent variable and shows that VTS shifts are 0.9 percent longer. Combining the estimates from Models 1 and 4 yields an implied labor-supply elasticity of approximately 1.8.

Table 3 reports the OLS and IV estimates of the labor-supply elasticity. Column 1 presents the OLS estimate, which reflects the correlation between hourly wages and hours worked after controlling for proportion of trips paid by credit card, share of trips in each borough, driver, start date and start hour fixed effect. The estimate implies that a 1 percent increase in hourly wages is associated with a 0.267 percent reduction in hours worked. Column 2 presents the instrumental-variables estimate, which isolates exogenous wage variation induced by taximeter type. The IV coefficient is positive and statistically significant: a 1 percent increase in hourly wages leads to a 1.7 percent increase in hours worked. This estimate represents the labor-supply elasticity and indicates that, when drivers face higher hourly earnings, they respond by working substantially longer shifts rather than reducing hours or ending their shift early.

This elasticity of roughly 1.7 is consistent with the neoclassical model of labor supply, which predicts that workers increase hours when wages rise. It implies that New York City yellow-cab drivers increase their work hours more than proportionally when their hourly wage increases. Although the estimate is larger than typical labor-supply elasticities found in other settings, often between 0.1 and 0.5, it is reasonable for two key reasons. First, our estimate reflects short-run adjustments rather than long-run labor-market decisions such as labor-force participation or full-time versus part-time work. Short-run elasticities can be substantially larger because drivers can adjust their hours immediately in response to higher pay. Second, taxi drivers have an unusual degree of temporal flexibility: they can decide at any moment whether to continue working or end their shift, unlike most workers who face fixed schedules or employer-determined hours. This flexibility allows their labor supply to respond more strongly to short-term wage variation.

7 Robustness checks

We conduct two robustness checks to assess whether our results depend on the specific sample definition or on the inclusion of control variables. First, Table 4 re-estimates the elasticity under progressively richer specifications. Model 1, which omits all controls and fixed effects, yields a negative elasticity. Model 2 adds the control vector and leaves the estimate essentially unchanged, so the negative sign is not an artifact of payment composition or borough mix. The elasticity turns positive only in Model 3, once driver, date, and hour-of-day fixed effects are absorbed. This indicates that the relevant confounders are fixed driver attributes and common time-varying conditions rather than the composition of trips within a shift. This importance of driver fixed effect in estimating elasticity is also highlighted in Motghare (2021).

Second, we vary the time gap used to delimit a shift and re-estimate the elasticity under each alternative definition in Table 5. Because a shift is not directly observed and must be inferred from the spacing of trips, the six-hour threshold used throughout is necessarily a judgment call, and a finding that survives only at that threshold would be difficult to interpret. The estimates move little across the alternatives, both in magnitude and in precision, so the elasticity we report is a feature of driver behavior rather than of the shift-construction rule. Taken together, the two exercises locate the identifying variation squarely within drivers and across shifts, which is the variation the model requires.

8 Heterogeneity

We examine heterogeneity of the elasticity estimate along three dimensions: medallion ownership, driver experience, and time of day. Columns 1 and 2 of Table 6 estimate the elasticity separately for owner-drivers and renter-drivers. For owner-drivers, the first-stage effect of the instrument on hourly wage is positive (0.016) and significant, but the F-statistic is only 4.94, much smaller than 10, the rule-of-thumb proposed by Staiger and Stock (1997). Because owners rarely switch taximeters, very little wage variation remains that can be attributed to the instrument. As a result, the owner sample does not contain independent wage variation sufficient to identify the elasticity, and we do not interpret the resulting estimate (–0.293) as elasticity of owners. This issue does not arise for renter-drivers, who exhibit a strong first-stage relationship between meter type and hourly wage. Their elasticity estimate (1.34) is well-identified and closely aligned with the overall elasticity estimate.

Columns 3 and 4 of Table 6 estimate the elasticity separately for drivers who primarily work day shifts and night shifts. For day-shift drivers, the first-stage effect of the instrument on hourly wage is positive (0.004) and statistically significant, but the F-statistic is only 9.24. Because this falls below conventional thresholds, we do not interpret the resulting elasticity estimate (3.881). Night-shift drivers do not face a weak-instrument problem: the instrument increases their hourly wage by 0.9 percent, yet they do not adjust their hours in response, yielding an elasticity estimate of 0.069 that is statistically indistinguishable from zero. The time of day explains this pattern. A day-shift driver ending a shift in the late afternoon can continue working into the evening rush, whereas a night-shift driver finishing in the early morning cannot—leases expire, demand thins, and fatigue binds. Our elasticity is measured at an interior margin, and the night shift simply does not offer one.

Columns 2–4 of Table 7 report elasticity estimates for drivers with different levels of experience, proxied by the number of shifts completed in 2010. For drivers in the first tercile (those with the fewest shifts), the first stage is strong and the elasticity estimate (1.4) is smaller than the overall estimate. For drivers in the second tercile, the elasticity estimate is slightly higher and close to the overall estimate. For drivers in the third tercile, however, the first stage is weak ($F = 2.82$), and the resulting elasticity estimate is not interpretable. Thus, elasticity appears to increase from the lowest tercile to the middle tercile, consistent with French and Stafford (2026), who finds that labor supply is more elastic among more experienced workers.

To summarize, we cannot identify elasticity for owner-drivers, day-shift drivers, or the most experienced drivers because meter type generates too little independent wage variation, resulting in a weak-instrument problem. Among the well-identified subsamples, renter-drivers exhibit an elasticity similar to the overall estimate, night-shift drivers show an elasticity close to zero, and elasticity rises with experience from the lowest to the middle tercile. Overall, the subsample results provide no evidence of negative elasticities.

9 Discussion

Our estimated elasticity of 1.7 is large relative to the broader labor-supply literature, but it is consistent with evidence for workers who have substantial discretion over their hours. The 95-percent confidence interval, [0.58, 2.79], excludes several prominent benchmarks: the Congressional Budget Office's central estimate of 0.40 (Whalen and Reichling, 2017), the intensive-margin Frisch elasticity of 0.54 reported by Chetty et al. (2011), and the meta-analytically corrected range of roughly 0.2 to 0.25 (Elminejad et al., 2023). At the same time, the interval encompasses Ham and Reilly (2013) Frisch elasticity of approximately 1.0 for workers operating under implicit contracts and lies close to the 90th-percentile elasticity of 1.92 reported by Attanasio et al. (2018) for the most responsive workers in their sample. Both of these point estimates fall within one standard error of ours.

Within the taxi-driver literature, our findings align with studies showing positive wage elasticities—Farber (2005, 2008, 2015) and Fehr and Goette (2007)—rather than the negative elasticities reported by Camerer et al. (1997) and Chou (2002), which were interpreted as evidence of daily income targeting. Our contribution adds a new source of plausibly exogenous wage variation:

taximeter vendor. Because meter type affects only the default tip prompts and is orthogonal to demand shocks and division bias, it provides cleaner identification than earlier approaches on both sides of the debate.

Taken together, the evidence suggests that our estimate reflects the upper end of a genuinely heterogeneous elasticity distribution concentrated among workers who set their hours in real time. Short-run, flexible-margin environments like taxi driving naturally yield higher elasticities than the population average implied by salaried workers with fixed schedules, consistent with Attanasio et al. (2018) argument that the aggregate elasticity is not a single structural parameter.

The methodological advantages of our setting come at the cost of generalizability. New York City yellow-cab drivers are a highly specific group and do not represent the broader labor market. Nonetheless, transparent identification for a well-defined subsample improves our understanding of labor-supply behavior and cautions against applying aggregate elasticities to contexts where workers face very different constraints on their hours.

10 Conclusion

This study re-examines how New York City taxi drivers respond to changes in their hourly wages using a new instrumental identification strategy based on differences in default tip suggestions across taximeters. We find that high-default taximeters increase drivers' credit card tip earnings by 7%, translating to a 0.5% increase in total hourly wages. This wage increase causes drivers to work longer shifts: a 1% increase in hourly wages leads to a 1.7% increase in hours worked. The result is robust to using different gaps to define shifts and model specifications.

Our findings support the standard economic prediction that people work more when wages are higher, contrasting with earlier research suggesting taxi drivers behave as "income targeters." They contribute to understanding how workers respond to temporary wage changes when they have schedule flexibility. This has relevance for the growing gig economy, where workers increasingly control when and how much they work. It also informs debates about overtime pay and flexible work arrangements—our results suggest that combining higher pay with schedule flexibility can lead to substantial increases in labor supply. Several limitations point to directions for future research. We focus on a single year (2010) in a single city, we cannot observe longer-term entry and exit decisions.

The methodological approach we develop—using institutional variation in payment technology as an instrument for wages—provides a template for studying labor supply in other flexible work contexts. Future work could examine more recent years as the taxi industry faces competition from ride-sharing services, compare results across multiple cities, analyze how different types of drivers respond to wage changes, and incorporate data on entry and exit decisions to estimate longer-run elasticities.

References

Angrist, J. D., S. Caldwell, and J. V. Hall (2021). Uber versus taxi: A driver's eye view. *American Economic Journal: Applied Economics 13*(3), 272–308.

Attanasio, O., P. Levell, H. Low, and V. Sánchez-Marcos (2018). Aggregating elasticities: intensive and extensive margins of women's labor supply. *Econometrica 86*(6), 2049–2082.

Bailey, M. J., J. DiNardo, and B. A. Stuart (2021, April). The Economic Impact of a High National Minimum Wage: Evidence from the 1966 Fair Labor Standards Act. *Journal of Labor Economics 39*(S2), S329–S367.

Barlevy, G. and D. Neal (2019, January). Allocating Effort and Talent in Professional Labor Markets. *Journal of Labor Economics 37*(1), 187–246.

Borjas, G. J. (1980). The relationship between wages and weekly hours of work: The role of division bias. *The Journal of Human Resources 15*(3), 409–423.

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

Chetty, R., A. Guren, D. Manoli, and A. Weber (2011). Are micro and macro labor supply elasticities consistent? a review of evidence on the intensive and extensive margins. *American Economic Review 101*(3), 471–475.

Chou, Y. K. (2002). Testing alternative models of labour supply: Evidence from taxi drivers in singapore. *The Singapore Economic Review 47*(01), 17–47.

Cortes, P. and J. Pan (2013). Outsourcing household production: Foreign domestic workers and native labor supply in hong kong. *Journal of Labor Economics 31*(2), 327–371.

Cortés, P. and J. Pan (2019). When time binds: Substitutes for household production, returns to working long hours, and the skilled gender wage gap. *Journal of Labor Economics 37*(2), 351–398.

Cortés, P. and J. Tessada (2011, July). Low-Skilled Immigration and the Labor Supply of Highly Skilled Women. *American Economic Journal: Applied Economics 3*(3), 88–123.

Eissa, N. and H. W. Hoynes (2004, August). Taxes and the labor market participation of married couples: the earned income tax credit. *Journal of Public Economics 88*(9), 1931–1958.

Eissa, N. and H. W. Hoynes (2006, January). Behavioral Responses to Taxes: Lessons from the EITC and Labor Supply. *Tax Policy and the Economy 20*, 73–110.

Elminejad, A., T. Havranek, R. Horvath, and Z. Irsova (2023). Intertemporal substitution in labor supply: A meta-analysis. *Review of Economic Dynamics 51*, 1095–1113.

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

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

Farber, H. S. (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 L. Goette (2007). Do workers work more if wages are high? evidence from a randomized field experiment. *American Economic Review 97*(1), 298–317.

French, E. and J. Jones (2012, April). Public pensions and labor supply over the life cycle. *International Tax and Public Finance 19*(2), 268–287.

French, S. and T. Stafford (2026). Returns to experience and the elasticity of labor supply. *Review of Economic Dynamics*, 101323.

Gai, Q., N. Guo, B. Li, Q. Shi, and X. Zhu (2025). Rural Pensions, Labor Reallocation, and Aggregate Income: An Empirical and Quantitative Analysis of China. *Econometrica 93*(5), 1663–1696.

Giuliano, L. (2013). Minimum Wage Effects on Employment, Substitution, and the Teenage Labor Supply: Evidence from Personnel Data. *Journal of Labor Economics 31*(1), 155–194.

Haggag, K. and G. Paci (2014). Default tips. *American Economic Journal: Applied Economics 6*(3), 1–19.

Ham, J. C. and K. T. Reilly (2013). Implicit contracts, life cycle labor supply, and intertemporal substitution. *International Economic Review 54*(4), 1133–1158.

Hoynes, H. W. (2000). Local Labor Markets and Welfare Spells: Do Demand Conditions Matter? *The Review of Economics and Statistics 82*(3), 351–368.

Hunt, J. (1999). Has work-sharing worked in germany? *The Quarterly journal of economics 114*(1), 117–148.

Hunt, J. (2012). Flexible Work Time in Germany: Do Workers Like It and How Have Employers Exploited It over the Cycle? *SOEPpapers on Multidisciplinary Panel Data Research*. Number: 489.

Kuhn, P. and F. Lozano (2008, April). The Expanding Workweek? Understanding Trends in Long Work Hours among U.S. Men, 1979–2006. *Journal of Labor Economics 26*(2), 311–343.

Kuhn, P. and C. Riddell (2010). The Long-Term Effects of Unemployment Insurance: Evidence from New Brunswick and Maine, 1940â€"1991. *ILR Review 63*(2), 183–204.

Lamadon, T., J. Lise, C. Meghir, and J.-M. Robin (2024). Labor market matching, wages, and amenities. Technical report, National Bureau of Economic Research.

Lamadon, T., M. Mogstad, and B. Setzler (2022, January). Imperfect Competition, Compensating Differentials, and Rent Sharing in the US Labor Market. *American Economic Review 112*(1), 169–212.

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

Martínez, I. Z., E. Saez, and M. Siegenthaler (2021, February). Intertemporal Labor Supply Substitution? Evidence from the Swiss Income Tax Holidays. *American Economic Review 111*(2), 506–546.

Motghare, S. (2021). The long-run elasticity of labor supply: New evidence for new york city taxicab drivers. *Labour Economics 71*, 102025.

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

Prendergast, C. (2002). The Tenuous Trade-off between Risk and Incentives. *Journal of Political Economy 110*(5), 1071–1102.

Siegenthaler, M. (2015). Has Switzerland Really Been Marked by Low Productivity Growth? Hours Worked and Labor Productivity in Switzerland in a Long-run Perspective. *Review of Income and Wealth 61*(2), 353–372.

Stafford, T. M. (2015). What do fishermen tell us that taxi drivers do not? an empirical investigation of labor supply. *Journal of Labor Economics 33*(3), 683–710.

Stafford, T. M. (2018). Do workers work more when earnings are high? *IZA World of Labor*, 1–10.

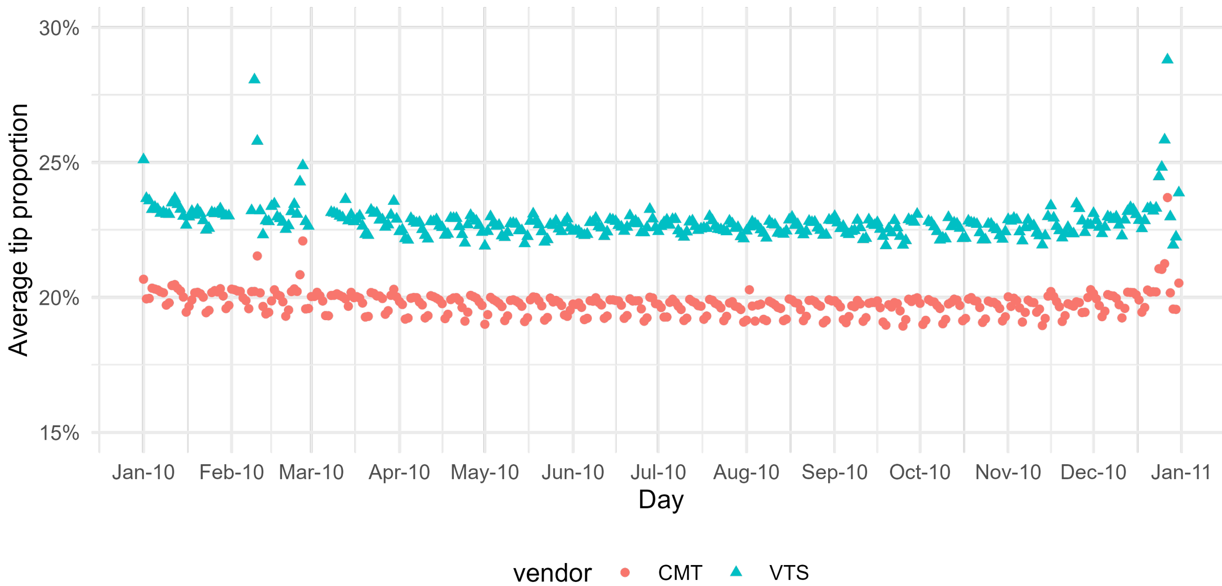
Staiger, D. and J. H. Stock (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica 65*(3), 557–586.

Stoddard, C. and P. Kuhn (2004, November). Incentives and Effort in the Public Sector: Have U.S. Education Reforms Increased Teachers' Work Hours?

Whalen, C. and F. Reichling (2017). Estimates of the frisch elasticity of labor supply: A review. *Eastern Economic Journal 43*(1), 37–42.

Figures/Tables

Figure 1: Average Tip Proportion



Notes: This figure presents the average ratio of tip amount to fare amount across all trips by taximeter type. The ratio is higher for trips completed using VTS taximeters than for trips completed using CMT taximeters. Because tips account for approximately 20–25 percent of total fare revenue, the difference in tip-to-fare ratios indicates that taximeter type is systematically associated with tip earnings. This pattern provides descriptive evidence that the VTS indicator generates variation in hourly wages through differences in passenger tipping behavior.

Table 1: Summary Statistics by Vendor: Year 2010

	CMT				VTS			
	N	**Mean**	**Min**	**Max**	**N**	**Mean**	**Min**	**Max**
Renter	2,456,660	0.93	0.00	1.00	2,548,182	0.87	0.00	1.00
Shift duration	2,913,922	9.41	1.48	24.00	3,208,203	9.38	1.57	23.98
Hourly wage								
Including tips	2,913,922	28.05	5.05	50.00	3,208,203	28.33	5.11	50.00
Excluding tips	2,913,922	26.11	4.64	49.93	3,208,203	26.19	5.02	49.38
Tip wage	2,913,922	1.95	0.00	25.60	3,208,203	2.14	0.00	28.10
Card tips	2,913,922	1.94	0.00	25.60	3,208,203	2.14	0.00	28.10
Credit trips prop.	2,913,922	0.36	0.00	1.00	3,208,203	0.37	0.00	1.00
Airport trips								
JFK	2,913,922	0.01	0.00	0.91	3,208,203	0.01	0.00	1.00
Newark	2,913,922	0.00	0.00	0.36	3,208,203	0.00	0.00	0.31
Pick-up location								
Staten Island	2,913,922	0.00	0.00	1.00	3,208,203	0.00	0.00	1.00
Manhattan	2,913,922	0.92	0.00	1.00	3,208,203	0.89	0.00	1.00
Bronx	2,913,922	0.00	0.00	1.00	3,208,203	0.00	0.00	1.00
Brooklyn	2,913,922	0.02	0.00	1.00	3,208,203	0.02	0.00	1.00
Queens	2,913,922	0.04	0.00	1.00	3,208,203	0.05	0.00	1.00

Red = CMT mean; Blue = VTS mean.

Notes: This table reports summary statistics for all twelve months sample in 2010 separately for shifts completed using VTS and CMT taximeters. The unit of observation is a driver–shift. Hours worked measure the total duration of a driver's shift. Hourly wage is calculated as total revenue earned during the shift divided by hours worked and includes both fare and tip revenue. Fare wage and tip wage represent fare revenue per hour and tip revenue per hour, respectively. The share of credit-card trips measures the fraction of trips within a shift paid electronically. Borough shares report the proportion of trips originating in Manhattan, the Bronx, Brooklyn, and Queens, with Staten Island omitted as the reference category. The final column reports the difference in means between VTS and CMT shifts. During the study period, both taximeter systems calculated fares using identical fare schedules but differed only in the default tip suggestions displayed to passengers. Consequently, systematic differences are expected primarily in tip-related outcomes rather than fare-related outcomes. The similarity of observable shift characteristics across vendors provides preliminary evidence supporting the identifying assumption that taximeter type affects driver behavior only through its influence on passenger tipping behavior.

Table 2: First Stage, Fare Wage, Card Tip, and Reduced Form Estimates: 2010

	First Stage ln(Wage) (1)	Fare Wage ln(Fare Wage) (2)	Card Tip ln(Card Tip Wage) (3)	Reduced Form ln(Shift Duration) (4)
VTS	0.005***	0.000	0.074***	0.009***
	(0.001)	(0.001)	(0.001)	(0.002)
Observations	5,656,890	5,656,890	5,639,693	5,656,890
First-stage F-statistic	31.99			

Notes: Controls: credit trip share; borough trip shares (Manhattan, Bronx, Brooklyn, Queens).
All specifications include driver, start date, and start hour fixed effects. SE clustered by driver.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Notes: This table reports the first-stage estimates corresponding to Equation (13). The dependent variable is the natural logarithm of hourly wage, fare wage, or tip wage, as indicated in each column. The key explanatory variable is an indicator equal to one if the taxi is equipped with a VTS taximeter and zero if it is equipped with a CMT taximeter. All specifications include driver fixed effects, calendar-date fixed effects, hour-of-day fixed effects, and the baseline control variables consisting of the share of credit-card trips and borough trip shares. Standard errors, reported in parentheses, are clustered at the driver level. The first-stage F-statistic is reported to assess instrument relevance. Because the two taximeter systems calculate fares identically but differ in the default tip prompts shown to passengers, the instrument is expected to affect tip wage but not fare wage, providing additional support for the proposed identification strategy.

Table 3: OLS and IV Estimates of Intertemporal Labor Supply: 2010

	OLS (1)	IV (2)
ln(Wage)	-0.267***	1.685***
	(0.002)	(0.564)
Observations	5,656,890	5,656,890
Kleibergen-Paap Wald F		31.99

Notes: Dependent variable is ln(Shift Duration). Controls: credit trip share;
borough trip shares (Manhattan, Bronx, Brooklyn, Queens). All specifications include
driver, start date, and start hour fixed effects. SE clustered by driver.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Notes: The dependent variable is the natural logarithm of hours worked during a shift. Ordinary Least Squares (OLS) estimates correspond to Equation (12), while instrumental-variables (IV) estimates correspond to Equations (13) and (14). Hourly wage is instrumented using an indicator equal to one if the taxi is equipped with a VTS taximeter. All specifications include driver fixed effects, calendar-date fixed effects, hour-of-day fixed effects, and baseline control variables consisting of the share of credit-card trips and borough trip shares. Standard errors, reported in parentheses, are clustered at the driver level. The IV estimates identify the causal effect of hourly wages on hours worked using only exogenous wage variation generated by differences in default tip prompts across taximeter vendors.

Table 4: Robustness: Varying Controls and Fixed Effects, 2010

	No Controls, No FEs (1)	Controls Only, No FEs (2)	Controls and FEs (3)
Panel A: First Stage — DV = ln(Wage)			
VTS	0.011***	0.007***	0.005***
	(0.001)	(0.001)	(0.001)
Observations	5,657,430	5,657,430	5,656,890
First-stage F-statistic	86.32	31.65	31.99
Panel B: Reduced Form — DV = ln(Shift Duration)			
VTS	-0.012***	-0.017***	0.009***
	(0.002)	(0.002)	(0.002)
Observations	5,657,430	5,657,430	5,656,890
Panel C: IV — DV = ln(Shift Duration)			
ln(Wage)	-1.072***	-2.469***	1.685***
	(0.164)	(0.449)	(0.564)
Observations	5,657,430	5,657,430	5,656,890
Kleibergen-Paap Wald F	86.32	31.65	31.99
Controls	No	Yes	Yes
Driver FE	No	No	Yes
Date FE	No	No	Yes
Hour FE	No	No	Yes

Notes: 2010 sample. Controls: credit trip share; borough trip shares (Manhattan, Bronx, Brooklyn, Queens, JFK, Newark). SE clustered by driver in all columns. Column (3) is the baseline specification. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Notes: This table examines the robustness of the baseline instrumental-variables estimates to alternative sets of control variables. Each column modifies the baseline specification by adding or removing observed controls while maintaining the same instrumental-variables strategy. Unless otherwise noted, all specifications include driver fixed effects, calendar-date fixed effects, and hour-of-day fixed effects, and hourly wage is instrumented using the VTS taximeter indicator. Standard errors are clustered at the driver level.

Table 5: OLS and IV Estimates of Intertemporal Labor Supply by Shift-Gap Threshold: 2010

	OLS (1)	IV (2)
Panel A: 5-Hour Shift Gap		
ln(Wage)	-0.268***	1.466***
	(0.002)	(0.491)
Observations	5,671,714	5,671,714
Kleibergen-Paap Wald F		38.54
Panel B: 6-Hour Shift Gap		
ln(Wage)	-0.267***	1.685***
	(0.002)	(0.564)
Observations	5,656,890	5,656,890
Kleibergen-Paap Wald F		31.99
Panel C: 7-Hour Shift Gap		
ln(Wage)	-0.269***	1.893***
	(0.002)	(0.640)
Observations	5,634,884	5,634,884
Kleibergen-Paap Wald F		26.96
Panel D: 8-Hour Shift Gap		
ln(Wage)	-0.272***	1.964***
	(0.002)	(0.673)
Observations	5,599,647	5,599,647
Kleibergen-Paap Wald F		24.89

Notes: Dependent variable is ln(Shift Duration). Each panel uses shifts constructed under a different minimum gap threshold (Panel A: 5 hours; Panel B: 6 hours; Panel C: 7 hours; Panel D: 8 hours). Controls: credit trip share; borough trip shares (Manhattan, Bronx, Brooklyn, Queens). All specifications include driver, start date, and start hour fixed effects. SE clustered by driver.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Notes: This table reports OLS and instrumental-variables estimates using alternative sample restrictions based on minimum shift-length thresholds. Each column re-estimates the baseline specification after applying the indicated threshold for hours worked. The dependent variable is the natural logarithm of hours worked. In the IV specifications, hourly wage is instrumented using the VTS taximeter indicator. All regressions include the baseline control variables, driver fixed effects, calendar-date fixed effects, and hour-of-day fixed effects. Standard errors are clustered at the driver level. Comparing estimates across alternative thresholds evaluates whether the estimated labor supply elasticity is sensitive to the choice of sample restriction.

Table 6: Heterogeneity by Medallion Ownership and Time of Shift, 2010

	Medallion Ownership		Time of Shift	
	Owner	Renter	Day Shift	Night Shift
	(1)	(2)	(3)	(4)
Panel A: First Stage — DV = ln(Wage)				
VTS	0.016**	0.006***	0.004***	0.009***
	(0.007)	(0.001)	(0.001)	(0.002)
Observations	449,757	4,171,247	2,188,363	2,232,350
First-stage F-statistic	4.94	36.98	9.24	32.03
Panel B: Reduced Form — DV = ln(Shift Duration)				
VTS	-0.005	0.009***	0.016***	0.001
	(0.017)	(0.002)	(0.003)	(0.002)
Observations	449,757	4,171,247	2,188,363	2,232,350
Panel C: IV — DV = ln(Shift Duration)				
ln(Wage)	-0.293	1.338***	3.881**	0.069
	(0.999)	(0.471)	(1.602)	(0.233)
Observations	449,757	4,171,247	2,188,363	2,232,350
Kleibergen-Paap Wald F	4.94	36.98	9.24	32.03

Notes: Controls: credit trip share; borough trip shares (Manhattan, Bronx, Brooklyn, Queens). All specifications include driver, start date, and start hour fixed effects. SE clustered by driver. Day shift: >75% of shifts start 4am–10am. Night shift: >75% of shifts start 2pm–8pm.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Notes: This table reports instrumental-variables estimates separately by rental status and shift type. Drivers are classified according to whether they rent or own the taxi medallion and whether they primarily operate during the day or at night. The dependent variable is the natural logarithm of hours worked, and hourly wage is instrumented using the VTS taximeter indicator. All specifications include the baseline control variables, driver fixed effects, calendar-date fixed effects, and hour-of-day fixed effects. Standard errors are clustered at the driver level. The estimates evaluate whether labor supply responses to exogenous wage changes differ across ownership arrangements and work schedules.

Table 7: Heterogeneity by Driver Experience (Shift Terciles), 2010

	Full Sample (1)	1st Tercile (Fewest Shifts) (2)	2nd Tercile (3)	3rd Tercile (Most Shifts) (4)
Panel A: First Stage — DV = ln(Wage)				
VTS	0.005***	0.011***	0.006***	0.002*
	(0.001)	(0.002)	(0.001)	(0.001)
Observations	5,656,890	623,443	2,046,248	2,987,199
First-stage F-statistic	31.99	23.20	17.62	2.82
Panel B: Reduced Form — DV = ln(Shift Duration)				
VTS	0.009***	0.015***	0.010***	0.004
	(0.002)	(0.005)	(0.003)	(0.003)
Observations	5,656,890	623,443	2,046,248	2,987,199
Panel C: IV — DV = ln(Shift Duration)				
ln(Wage)	1.685***	1.437**	1.664**	1.809
	(0.564)	(0.589)	(0.752)	(2.021)
Observations	5,656,890	623,443	2,046,248	2,987,199
Kleibergen-Paap Wald F	31.99	23.20	17.62	2.82

Notes: Controls: credit trip share; borough trip shares (Manhattan, Bronx, Brooklyn, Queens). All specifications include driver, start date, and start hour fixed effects. SE clustered by driver. Terciles defined by total number of shifts per driver in 2010. Column (1) is the full sample.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Notes: This table reports instrumental-variables estimates separately for terciles of driver experience. Drivers are classified into low-, middle-, and high-experience groups according to the distribution of observed driving experience in the sample. The dependent variable is the natural logarithm of hours worked, and hourly wage is instrumented using the VTS taximeter indicator. All specifications include the baseline control variables, driver fixed effects, calendar-date fixed effects, and hour-of-day fixed effects. Standard errors are clustered at the driver level. The estimates evaluate whether labor supply responses to exogenous wage changes differ across drivers with different levels of experience.

A Data Appendix

A.1 Computing Work Hours and Hourly Wage from Trip Data

We use complete trip records for all NYC yellow cabs from January through December 2010. For each trip, the data report the encrypted hack ID (driver license number), medallion ID, trip start and end times, fare amount, taxes, surcharges, tolls, payment method, number of passengers, trip distance, trip duration, and tip amount (for credit-card payments). Because the encrypted driver license number is constant throughout the year, we observe every trip made by each driver during the study period.

Each month contains two files—a trip file and a fare file—with one observation per trip. The trip file includes pickup and drop-off locations, trip distance, and trip time; the fare file includes fare details. We merge these two files for each month to create a combined trip-level dataset.

Trip-Level Cleaning

We apply several consistency checks and drop trips that satisfy any of the following criteria:

- Trip start time occurs after the trip end time.
- Trip duration (based on timestamps) is less than 10 seconds.[5]
- Fare amount is zero or exceeds $10,000, indicating abnormal or erroneous trips.
- Distance traveled is non-positive or greater than 1,000 miles.
- Payment type is "no charge," "dispute," "unknown," or "voided trip"; we retain only cash and credit-card trips.
- Duplicate trips with identical driver–medallion–start-time combinations.
- Trips where the end time of one trip is later than the start time of the next trip for the same driver.

Constructing Shifts

Trip durations alone do not capture total work hours because they exclude time spent searching for passengers. To approximate actual work time, we aggregate trips into shifts, defined as sequences of trips separated by less than six hours. For each shift, we compute:

- **Hours worked**: difference between the start time of the first trip and the end time of the last trip.
- **Shift revenue**: sum of fare amounts across all trips in the shift.

[5] We use duration based on trip start and end times over the already defined trip duration variable because for quite a few observations, the recorded trip duration seems implausible.

- Additional shift-level variables used in Table 1.

This procedure yields a shift-level dataset where each observation represents a continuous work period.

Shift-Level Cleaning

We drop shifts that satisfy any of the following criteria:

- Total hours worked less than 1 hour or greater than 24 hours.
- Fewer than 5 trips in the shift.
- Hourly wage below $5 or above $50, to reduce the influence of outliers.

B Identifying Owner-Drivers and Renter-Drivers

The raw data contain encrypted medallion and driver license numbers. The encryption method is MD5.[6] We download the publicly available medallion and driver license numbers from NYC OpenData and generate their MD5 hashes. Matching these hashes to the raw data allows us to recover the actual medallion and driver IDs.

To classify drivers as owners or renters, we use archived TLC lists of Medallion Leases, which report long-term lease arrangements.[7] These lists identify drivers who own medallions and the corresponding medallion numbers. We also use a February 2012 snapshot of medallion types and driving arrangements from the TLC. Owner-drivers are defined as drivers who appear on the long-term lease list without a lease end date for exactly one independent medallion and are never observed driving a mini-fleet medallion. Renters are defined as drivers who either drive mini-fleet medallions or appear as long-term drivers in the lease lists.

C Assigning Driver Types

We classify drivers along two dimensions: work intensity and shift timing.

- **Full-time drivers**: those who work at least 40 hours per week or complete at least five shifts per week.
- **Part-time drivers**: all others.

Shift timing is defined based on the start time of shifts:

- **Day-shift drivers**: more than 75% of shifts start between 4 a.m. and 10 a.m.

[6] `https://www.vijayp.ca/articles/blog/2014-06-21_on-taxis-and-rainbows--f6bc289679a1.html`
[7] Using `https://archive.org/` . Lists are dated December 06, 2012 and December 05, 2013.

- **Night-shift drivers**: more than 75% of shifts start between 2 p.m. and 8 p.m.
- **Other drivers**: those who do not meet either threshold.

D Assessing Endogenous Vendor Selection

Our instrumental variables strategy uses taximeter type (VTS versus CMT) as an instrument for hourly wages to estimate the labor supply elasticity of taxi drivers. The validity of this approach relies on the exclusion restriction: taximeter type must affect labor supply only through its effect on hourly wages and not through any other channel. A potential threat to this assumption would arise if drivers systematically selected which taximeter vendor to use based on expected earnings. In that case, taximeter type would become correlated with unobserved driver characteristics or preferences, violating the identifying assumptions of our IV strategy.

To evaluate this possibility, we examine the temporal stability of taximeter assignment at both the individual and aggregate levels. Our analysis focuses on drivers who worked continuously throughout the 11-month sample period in 2010. Restricting attention to balanced-panel drivers serves several purposes. First, it removes composition effects arising from drivers entering or leaving the market, allowing us to distinguish genuine changes in vendor assignment from changes in the sample. Second, these drivers had ample opportunity to learn about any persistent differences in earnings associated with the two taximeter systems. Third, because they were continuously active throughout the sample period, they represent the group most likely to optimize vendor choice if doing so were feasible. For each driver-month, we calculate the proportion of shifts completed using VTS and CMT taximeters.

We begin by examining individual driver behavior. Figure D1 through D4 present four representative patterns observed in the data: a driver who switches from VTS to CMT, a driver who consistently uses VTS throughout the sample period, a driver who switches from CMT to VTS, and a driver who consistently uses CMT throughout the sample period. These figures demonstrate that vendor assignment is heterogeneous across drivers. Some drivers remain with a single taximeter vendor throughout the sample, whereas others experience changes in assigned vendor over time. Importantly, however, switching occurs in both directions. Some drivers move from VTS to CMT, while others move from CMT to VTS. If drivers were actively selecting taximeters to maximize expected earnings, one would expect switching to occur predominantly toward VTS, since VTS displays higher default tip suggestions and therefore generates slightly higher hourly tip income on average. Instead, the observed switching patterns are bidirectional and do not exhibit any systematic movement toward the higher-paying vendor.

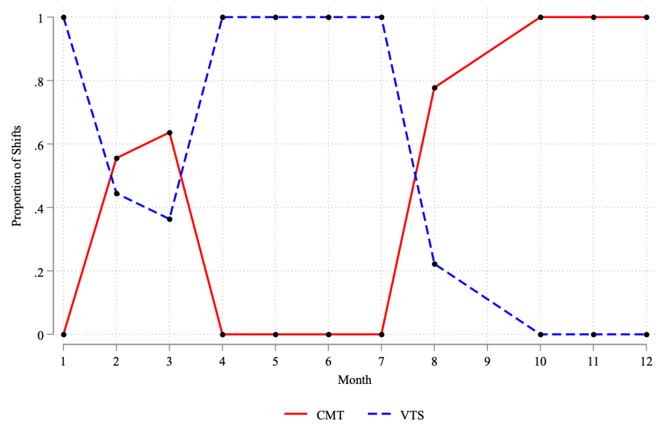
We next examine vendor assignment at the aggregate level. Figure D5 reports the monthly proportion of shifts completed using VTS and CMT taximeters across all balanced-panel drivers. Throughout the sample period, the market shares of the two vendors remain remarkably stable, with both systems accounting for approximately one-half of all observed shifts. Although modest month-to-month fluctuations are present, there is no persistent upward trend in VTS usage or corresponding decline in CMT usage. If drivers gradually learned about the small earnings advantage associated with VTS and were able to choose their taximeter vendor, one would expect VTS market share to increase over time. The absence of such a pattern suggests that vendor assignment is largely determined by institutional factors rather than driver optimization.

As a more demanding test, we repeat the analysis for a subset of highly active renter drivers who completed at least 20 shifts per month in at least 10 of the 11 sample months. These drivers represent the most experienced and professionally engaged participants in the market. They accumulated substantial exposure to both taximeter systems and, if vendor choice were feasible, would be the group most likely to recognize and exploit any systematic earnings differences. Figure D6 presents the monthly vendor shares for this restricted sample. The resulting pattern closely resembles that observed for the full balanced-panel sample. Vendor shares remain approximately evenly split between VTS and CMT throughout the sample period, with no sustained increase in VTS usage. Thus, even among the drivers with the greatest opportunity and strongest incentives to optimize vendor choice, we find no evidence of systematic movement toward the higher-paying taximeter system.

These findings provide evidence against endogenous vendor selection. The individual figures show that vendor switching occurs in both directions rather than systematically toward VTS, while the aggregate analyses reveal remarkably stable market shares over time for both the full sample of balanced-panel drivers and the subset of highly active renter drivers. Although these descriptive analyses cannot completely rule out occasional individual preferences or isolated instances of switching, they provide no evidence that drivers systematically select taximeter vendors based on expected earnings. Consequently, the observed patterns are consistent with the institutional setting in which renter drivers are assigned available taxis rather than freely choosing between vendors, lending additional support to the exclusion restriction underlying our instrumental variables strategy.

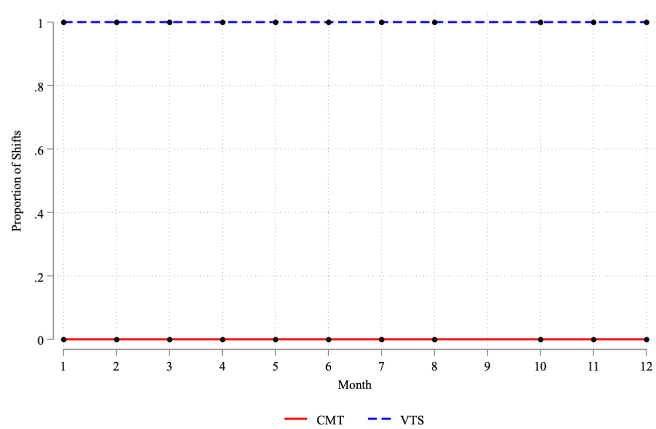
Appendix Figures

Figure D1: Representative Driver Switching from VTS to CMT



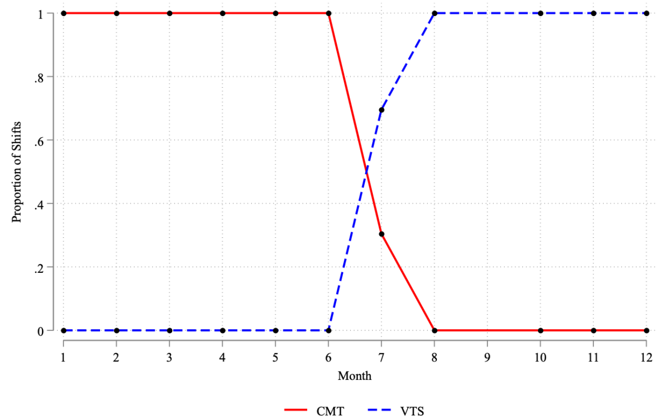
Note: This figure presents the monthly proportion of shifts completed using VTS and CMT taximeters for a representative driver who transitioned from predominantly using VTS-equipped taxis to predominantly using CMT-equipped taxis during the sample period. Each point represents the share of the driver's monthly shifts completed with each taximeter type. The gradual change in vendor usage demonstrates that some drivers experience changes in assigned taximeter type over time rather than remaining permanently attached to a single vendor.

Figure D2: Representative Driver with Persistent VTS Taximeter Use



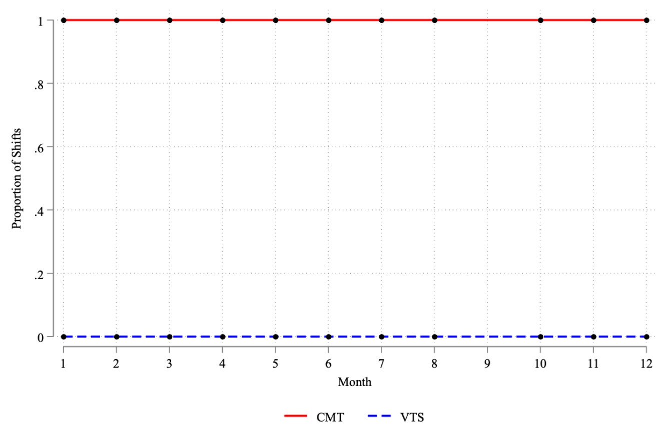
Note: This figure presents the monthly proportion of shifts completed using VTS and CMT taximeters for a representative driver who consistently operated a VTS-equipped taxi throughout the sample period. The figure shows that virtually all monthly shifts were completed using VTS taximeters, indicating no evidence of switching between taximeter vendors over time.

Figure D3: Representative Driver Switching from CMT to VTS



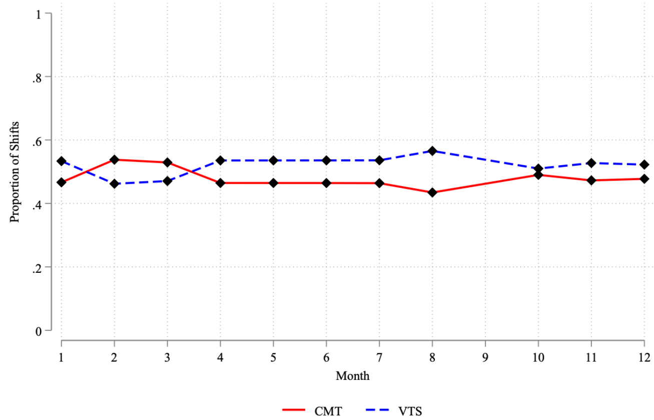
Note: This figure presents the monthly proportion of shifts completed using VTS and CMT taximeters for a representative driver who transitioned from predominantly using CMT-equipped taxis to predominantly using VTS-equipped taxis during the sample period. The observed change illustrates that vendor assignment can vary over time for some drivers, rather than remaining fixed throughout the sample.

Figure D4: Representative Driver with Persistent CMT Taximeter Use



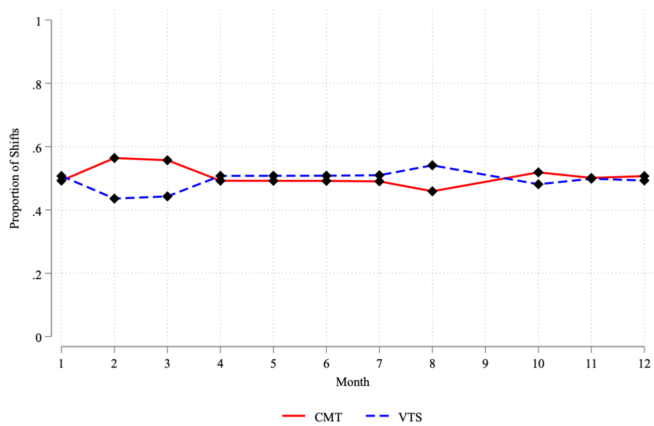
Note: This figure presents the monthly proportion of shifts completed using VTS and CMT taximeters for a representative driver who consistently operated a CMT-equipped taxi throughout the sample period. The figure shows that nearly all monthly shifts were completed using CMT taximeters, indicating no evidence of switching between taximeter vendors over time.

Figure D5: Vendor Choice Over Time Among Balanced Drivers



Note: This figure presents the monthly proportion of shifts completed using VTS and CMT taximeters among drivers observed in all eleven months of the sample period. Each month, we calculate the share of shifts associated with each taximeter type across all balanced drivers. The figure illustrates the stability of vendor assignment over time and shows that the market share of VTS and CMT taximeters remains approximately constant throughout the sample period.

Figure D6: Vendor Choice Over Time Among Frequent Renter Drivers



Note: This figure presents the monthly proportion of shifts completed using VTS and CMT taximeters among renter drivers who completed at least twenty shifts per month in at least ten months of the sample period. Restricting the sample to frequent renter drivers provides a setting where taxi assignment is less likely to reflect occasional participation decisions. The stability of VTS and CMT shares over time suggests that vendor assignment is driven primarily by institutional factors rather than systematic driver selection.