MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
https://pubsonline.informs.org/journal/msom
Vol. 25, No. 4, July–August 2023, pp. 1376–1393
ISSN 1523-4614 (print), ISSN 1526-5498 (online)

The Impact of Behavioral and Economic Drivers on Gig Economy Workers

Gad Allon,[a] Maxime C. Cohen,[b] Wichinpong Park Sinchaisri[c,*]

[a] The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104; [b] Desautels Faculty of Management, McGill University, Montreal, Quebec H3A 0G4, Canada; [c] Haas School of Business, University of California, Berkeley, California 94720

*Corresponding author

Contact: gadallon@wharton.upenn.edu, https://orcid.org/0000-0002-4815-6667 (GA); maxime.cohen@mcgill.ca, https://orcid.org/0000-0002-2474-3875 (MCC); parksinchaisri@haas.berkeley.edu, https://orcid.org/0000-0001-9351-0541 (WPS)

Received: October 28, 2019
Revised: June 1, 2021; July 14, 2022
Accepted: October 9, 2022
Published Online in Articles in Advance: March 2, 2023

https://doi.org/10.1287/msom.2023.1191

Abstract. ***Problem definition:*** Gig economy companies benefit from labor flexibility by hiring independent workers in response to real-time demand. However, workers' flexibility in their work schedule poses a great challenge in terms of planning and committing to a service capacity. Understanding what motivates gig economy workers is thus of great importance. In collaboration with a ride-hailing platform, we study how on-demand workers make labor decisions; specifically, whether to work and work duration. Our model revisits competing theories of labor supply regarding the impact of financial incentives and behavioral motives on labor decisions. We are interested in both improving how to predict the behavior of flexible workers and understanding how to design better incentives. ***Methodology/results:*** Using a large comprehensive data set, we develop an econometric model to analyze workers' labor decisions and responses to incentives while accounting for sample selection and endogeneity. We find that financial incentives have a significant positive influence on the decision to work and on the work duration—confirming the positive income elasticity posited by the standard income effect. We also find support for a behavioral theory as workers exhibit income-targeting behavior (working less when reaching an income goal) and inertia (working more after working for a longer period). ***Managerial implications:*** We demonstrate via numerical experiments that incentive optimization based on our insights can increase service capacity by 22% without incurring additional cost, or maintain the same capacity at a 30% lower cost. Ignoring behavioral factors could lead to understaffing by 10%–17% below the optimal capacity level. Lastly, our insights inform the design of platform strategy to manage flexible workers amidst an intensified competition among gig platforms.

Funding: This study was supported by The Jay H. Baker Retailing Center, The William and Phyllis Mack Institute for Innovation Management, The Wharton Risk Management and Decision Processes Center, and The Fishman-Davidson Center for Service and Operations Management.
Supplemental Material: The online appendices are available at https://doi.org/10.1287/msom.2023.1191.

Keywords: gig economy • labor supply • worker behavior • behavioral operations • empirical operations • incentives • sample selection

1. Introduction

Gig economy is a labor-sharing market system where workers engage in short-term projects or freelance work as opposed to permanent jobs. In 2021, 59 million Americans, or 36% of the U.S. workforce, engaged in gig work (Ozimek 2021), providing a wide range of services, from ride-hailing (e.g., Uber, Lyft) to food delivery (e.g., DoorDash, GrubHub) to web development (e.g., Upwork, Fiverr). The size of the independent workforce is growing three times faster than the overall U.S. workforce growth since 2014 and it is estimated that by 2025, the majority of the workforce will participate in the gig economy—leading to a global gross domestic product (GDP) boost of $2.7 trillion (Manyika et al. 2015). The unique and novel feature of this system relates to the nature of employment: independent workers can freely choose their work schedule as well as seamlessly switch between multiple platforms to provide service. Such flexibility attracts many workers to the gig economy.

Companies also greatly benefit from increased labor flexibility as they can hire workers with different skill levels to work at different times while compensating them for the work they perform. Like any other market, the key to success in the gig economy lies in the effective matching of supply with demand. Firms need to ensure that their services appeal not only to customers (demand) but also to independent service providers (supply). This poses an enormous challenge in planning and committing to a service capacity both during peak hours when demand is high and during off-peak times when only a handful of workers are needed. Policymakers have also joined the conversation, concerned with how such work structures might affect workers. For instance, New York City passed fatigued driving prevention rules as part of its Vision Zero initiative in 2017, limiting the number of daily and weekly hours a ride-hailing driver can work with the goal of reducing driver fatigue and enhancing road safety. In 2019, the European Parliament approved new rules that provide minimum rights and enforce better job transparency and compensation for gig workers.

To examine how firms can staff the right number of on-demand workers at the right time and how policymakers can develop effective regulations, it is important to first understand how gig workers make labor decisions. For decades, economists have studied how labor supply is influenced by economic incentives and behavioral motives. The standard income effect predicts that workers, as lifetime utility maximizers, are more likely to work or supply more labor in response to a higher wage. While several observational studies find evidence for this theory (e.g., Oettinger 1999, Sheldon 2016), other studies suggest the opposite prediction. New York City (NYC) taxi drivers are found to work for fewer hours on a high-paying day and more likely to quit working in response to higher accumulated income due to reference-dependent behavior with respect to earnings (e.g., Camerer et al. 1997, Thakral and Tô 2021). In other words, their decisions are based on reaching a target level of income or *income target*. Providing further support for the behavioral theory of labor supply, Crawford and Meng (2011) and Farber (2015) suggest that workers' behavior could perhaps be influenced by a target level of work duration or *time target*.

Our paper aims, in part, to reconcile this ongoing debate by proposing a framework to explain labor decisions through both economic incentives and behavioral motivations. Recent work in operations management in the context of the gig economy has focused on the system equilibrium or on social welfare (e.g., Cachon et al. 2017, Taylor 2018). To our knowledge, among the papers that focus on the supply side (e.g., Dong and Ibrahim 2020, Benjaafar et al. 2022), our work is the first to empirically examine the causal effect of behavioral and economic factors on gig economy workers' decisions and to incorporate their behavior into the optimization of financial incentives. Our work also follows calls for advancing behavioral operations research by studying worker behavior in new work environments such as on-demand services and freelancing platforms (Chen et al. 2020, Donohue et al. 2020).

1.1. Research Questions and Methodology

Our key research questions are (i) How do gig economy workers make labor decisions? How do they react to incentives? What factors shape their work schedule decisions? Are their decisions rational or do they exhibit behavioral biases? and (ii) How can gig companies set incentives to effectively recruit workers? How can they meet the desired service level by taking into account workers' behavior and offering them the right incentives?

We answer these questions by estimating an econometric model of workers' labor decisions and conducting numerical experiments on incentive optimization. Prior empirical studies on the relationship between wage and labor decisions have not distinguished between the decision of whether to work and the work duration decision and instead treated them essentially as a single decision due to data limitations. Through our collaboration with a U.S. ride-hailing company, we overcome this challenge by leveraging our rich data set, which contains real-time information on financial incentives regardless of drivers' subsequent labor decisions. Accordingly, we gain a clearer insight into drivers' decisions to work by investigating drivers who chose not to work during a particular period. In our empirical model, we address econometric challenges such as sample selection and omitted variable biases, and we account for drivers' heterogeneity and real-time market conditions and competition. Finally, we propose an optimization heuristic for targeted incentives and conduct counterfactual simulations to examine its performance and quantify potential losses if the company ignores workers' behavior when designing incentives.

1.2. Contributions

Our paper contributes to the economics and operations literatures in four ways. First, we are one of the first to offer a comprehensive empirical model to estimate workers' two-stage labor decisions, whether to work and work duration, conditional on financial incentives and personal targets, while taking into account sample selection, endogeneity, and market conditions. Gig economy workers' decisions have been empirically challenging to trace and study. Past studies have leveraged observational data of work activity and earnings from work performed by workers, and were focused on the work duration decision. Our unique data set allows us to develop a deeper understanding of the factors influencing the decision on whether to work as we observe both financial incentives offered by the focal platform and work decisions of all workers. Second, we offer a potential way to explain labor decisions by the two predominant theories of labor supply by showing that workers respond to wage variation in the same way as suggested by the standard income effect, while also exhibiting reference-dependent behavior with respect to accumulated earnings. We find that an hourly wage has a positive impact on both the decision to work and the work duration. However, our proxy for unobserved income targets—accumulated earnings from earlier hours of the same day or earlier days of the week—has a negative impact on both decisions. This finding provides support for an income-targeting behavior; that is, workers work less as they are closer to their income goal. In addition, we unravel a seemingly new behavioral phenomenon: our results indicate that workers' recent work duration (from earlier hours of the same day or earlier days of the same week) has a consistent and positive influence on the decision to continue working and on subsequent work duration. This phenomenon appears to capture the tendency of workers to make the same work decision as their recent ones. Third, we demonstrate that behavioral factors play an important role in workers' labor decisions. Both in-sample and out-of-sample analyses suggest that workers' reaction to accumulated earnings and past work duration are key drivers of their labor decisions. We then demonstrate via simulations that not accounting for these behavioral factors would result in understaffing by 10%–17%. Finally, we apply our insights to prescribe operational decisions and conduct regulatory impact analysis. Specifically, we show that if the company optimizes its incentive policy accounting for workers' behavior, it can increase its capacity by 22% without incurring additional cost or maintain the same service level at a 30% lower cost.

2. Labor Supply Theories and Hypotheses Development

Economists have offered two different perspectives centered around the elasticity of labor supply. On the one hand, the traditional approach follows a lifecycle model where individuals maximize their lifetime utility and predicts that workers exhibit positive income elasticity. On the other hand, empirical studies, notably in the context of taxi drivers, suggest that income elasticity could be negative if workers are loss averse and benchmark their earnings relative to a reference point. It is unclear whether existing findings can apply to gig economy workers who have full discretion over their work schedule. In this section, we review in greater detail the two contrasting models of labor supply and develop hypotheses for the behavior of gig economy workers.

2.1. Traditional Model of Labor Supply

In the neoclassical microeconomics tradition, each worker is a rational agent who maximizes lifetime utility. A positive wage shock should then lead to a larger group of workers joining the force or to a higher level of activity from workers. In other words, workers are expected to exhibit a positive wage elasticity (e.g., work more when facing a wage increase). This perspective seems plausible but finding evidence in the field has been challenging as in reality workers cannot easily adjust their work hours. However, positive elasticities have been observed among workers who have some level of discretion over their schedule, such as pipeline workers (Carrington 1996), vendors in a baseball stadium (Oettinger 1999), and fishermen (Stafford 2015). These studies find that wage shocks, typically driven by temporary demand variation, have a positive effect on labor supply—both on the number of workers and work hours.

2.2. Behavioral Model of Labor Supply

The seminal work by Camerer et al. (1997) studies NYC taxi drivers and finds substantial negative elasticities, suggesting that drivers' daily decisions on work hours are influenced by their individual income targets (known as the income-targeting effect). Using data from a different set of NYC taxi drivers, Farber (2005, 2008) finds that the probability to stop working is closely related to the realized income earned in the same day and it increases once the income target is reached, but concludes that the findings are not robust. Crawford and Meng (2011) implement similar econometric strategies to estimate models based on the reference-dependent preferences theory, which allows for consumption and gain-loss utilities. The authors conceptualize drivers' targeted levels of income and work hours and find that stopping probabilities are more influenced by the second target they reach on a given day. More recently, Thakral and Tô (2021) estimate a structural model of labor supply of NYC taxi drivers, allowing a time-dependent relationship between earnings and the stopping probability. Their results confirm that the income-targeting effect exists when controlling for the number of work hours. These findings offer a realistic behavioral explanation and align well with insights from behavioral economics; however, support for the behavioral theory has been lacking outside the taxi industry.

2.3. Labor Supply in the Gig Economy

The gig economy offers workers a flexible work schedule. As gig work appeals to a broad range of workers with different backgrounds and preferences, predicting the worker turnout or service capacity at any point in time is remarkably challenging. A common way to incentivize workers to join and to keep active workers engaged is to offer dynamic financial incentives. Real-time bonuses, such as Uber's surge prices and Caviar's Peak Pay, reward workers who work during busy periods with high demand. Beyond direct monetary rewards, several companies employ a combination of gamification and psychology and offer nonmonetary incentive programs. For example, Uber drivers can earn badges for achievements, from excellent service to entertaining ride, and are constantly reminded of how close they are to their earning goals. Whereas these incentive strategies are prevalent in practice, less is known in academic research about their influence on workers' labor decisions.

Our paper belongs to the fast-growing research trend that examines operational and pricing decisions in the context of the gig economy (for a review, see Benjaafar and Hu 2020). Most relevant to our work are studies that examine how dynamic wages affect supply and consider the problem of designing the optimal incentives to coordinate supply with demand for on-demand service platforms. Dynamic wages due to surge pricing have been shown to entice ride-hailing drivers to work longer (Chen 2016) and benefit drivers via better utilization (Cachon et al. 2017). Hu and Zhou (2020) study the contracts under which the platform takes a fixed cut from workers' earnings and demonstrates good performance among flat-commission contracts. Taylor (2018) shows that the uncertainty in workers' opportunity costs or in delay-sensitive customers' valuations can lead the intermediary to raise the price during congestion. Our work focuses on the supply side behavior and the need to use incentives to motivate flexible workers. There are relatively few studies that investigate worker behavior and its impact on the platform's operational decision. Most of these studies are of theoretical nature and focus on the equilibrium of matching supply with demand (see, e.g., Ibrahim 2018, Dong and Ibrahim 2020, Benjaafar et al. 2022).

The only empirical studies that incorporate worker behavior in a gig economy setting to our knowledge are Sheldon (2016), Karacaoglu et al. (2018), and Chen et al. (2019). Sheldon (2016) finds that Uber drivers' income elasticities are significantly positive and increasing over time, suggesting that if income targeting does exist, it would only be temporary and moderated by experience. Karacaoglu et al. (2018) study e-hailing taxi drivers in South America and find that drivers' response to real-time information about other drivers ' locations could explain different utilization they can achieve. Chen et al. (2019) document how Uber drivers value real-time flexibility and estimate the driver surplus from having a flexible schedule. The authors find that drivers earn higher surplus from Uber's flexible model relative to less flexible arrangements. Whereas these papers rigorously capture how gig workers respond to incentives and information, their models do not consider potential behavioral factors in explaining workers' behavior. This is due to data limitations given that most data sets record only the trips that happened. In our data set, however, we observe the information available to drivers even when they decided not to work. We focus on the behavior of gig workers and on how the platform can improve its operational decisions by understanding such behavior.

2.4. Hypotheses Development

We are interested in studying how gig economy workers make labor decisions, specifically whether they will work at a particular time and, if so, for how long. Labor decisions typically depend on multiple factors such as weather and external commitments. Yet, these are not controlled by the platform and, thus, while we attempt to control for such factors, we focus on the impact of economic drivers (hourly wage) and behavioral factors (workers' income and time targets). Several companies have exploited workers' tendency to set goals by helping workers track their progress toward the goals and nudging them to work for longer. Because individuals' targets cannot be observed, we use workers' accumulated earnings since the beginning of their workday as a proxy for their income target and the duration of their work so far as a proxy for their time target. We next present our hypotheses regarding the impact of each factor on gig economy workers' labor decisions.

Hypothesis 1. *A higher wage increases the probability of working and the work duration.*

Following the standard income effect (see Section 2.1), we expect that a higher hourly wage will increase the probability of working. Empirical studies of workers who have discretion over their work hours suggest that workers adjust labor decisions in the same direction as wage (see, e.g., Oettinger 1999, Stafford 2015). We posit that gig workers also exhibit a positive income elasticity as they have full control over their schedule. Unlike traditional employment, gig work tends to be smaller and temporary projects (e.g., assembling furniture, driving within a city) that require less time to complete. Consequently, work decisions are made more frequently and for a shorter time frame. The objective is therefore likely to maximize utility (e.g., earnings) in the following period. We still believe that there exists a behavioral explanation of labor supply, but such effect would be driven by accumulated earnings or work hours instead (see Hypotheses 2 and 3). Past studies that provide support for an income-targeting effect only modeled the relationship between the number of work hours and the average daily wage. We postulate that the negative impact on work duration will only be apparent during specific times of day (days of week), when workers might be closer to reaching their daily (weekly) income targets. Thus, when controlling for both accumulated income and work hours separately, we should observe a positive income elasticity.

Hypothesis 2. *Higher accumulated earnings decrease the probability of working and the work duration.*

Studies of taxi drivers, including Camerer et al. (1997), Farber (2008), and Thakral and Tô (2021), provide support for an income-targeting behavior; that is, the probability to stop working increases once the income target is reached. Thakral and Tô (2021) further demonstrate that drivers' decisions are highly influenced by recent earnings. Gig workers are also likely to be influenced by the income-targeting effect, as tracking their progress toward the income goal is much easier. Several gig platforms provide real-time information about workers' recent work activities and earnings through their apps and also provide frequent feedback about their earnings (e.g., after every completed trip for ride-hailing drivers). An alternative explanation of the negative impact of accumulated income is related to fatigue. Specifically, higher accumulated earnings could indicate a greater level of effort. Consequently, workers who experienced more fatigue would work for a shorter time. As a result, we expect to see a negative impact of the accumulated earnings on both the probability of working and the work duration.

Hypothesis 3. *Longer time worked decreases the probability of working and the work duration.*

Previous work in labor economics suggests another type of targeting behavior: time targeting. Crawford and Meng (2011) develop a structural stopping estimation model that allows for reference points in both daily income and work duration among taxi drivers and concludes that drivers are loss averse relative to both reference points. Agarwal et al. (2015) and Farber (2015) find that the probability of ending a work shift is positively related to cumulative work hours. As discussed in Hypothesis 2, fatigue could also be explained by work duration. Recent findings suggest that work performance deteriorates toward the end of long shifts among paramedics (Brachet et al. 2012) and part-time call center agents (Collewet and Sauermann 2017). Thus, we expect that the longer the workers have recently worked, the less likely they would continue working and, if they do work, the work duration would be shorter relative to those with a shorter past work duration.

3. Data: Ride-Hailing Platform in New York City

To answer our research questions, we collaborate with an on-demand ride-hailing company (referred to as "the company" or "the platform") and analyze a large comprehensive data set of driving activities and financial incentives in NYC over a period of 358 days (from October 2016 to September 2017). Our data includes each driver's vehicle type, experience with the platform, number of hours driven, and financial incentives offered and earned. The key advantage of our data is that we observe the incentives that were offered to every driver regardless of the decision to drive. In other words, even for drivers who decided not to drive for a particular time period, we still know their offered wage and promotions for that period. In total, we have several million driver-shift observations and several thousand unique drivers.[1] We next present an overview of the platform and report descriptive statistics of working shifts, financial incentives, and vehicle types.

3.1. Platform Overview

The company is a ride-hailing online platform that offers services in many cities worldwide. The users (riders) may request rides in real-time through a smartphone app. The platform then matches riders with available drivers. This platform offers a shared service (i.e., several passengers heading in the same direction may share the same vehicle). To make the service more efficient, passengers can be picked up and dropped off at an optimized location near the exact requested locations. Finally, the vast majority of drivers are compensated according to a guaranteed hourly rate regardless of the number of completed rides. We focus on drivers who are paid by the hour as this scheme resembles the traditional wage model but with more flexibility on the drivers' side. This allows us to investigate how drivers' work decisions are influenced by variations in monetary incentives.

3.2. Shifts and Work Schedule

Each operating day is divided into six shifts specified by the company (see an illustration in Figure 1): morning nonrush hours from midnight to 7 a.m. (AM off-peak), morning rush hours from 7 to 9 a.m. (AM peak), midday from 9 a.m. to 5 p.m. (midday), afternoon rush hours from 5 to 8 p.m. (PM peak), evening nonrush hours from 8 to 9 p.m. (PM off-peak), and late night from 9 p.m. to midnight (late night). The largest number of trips happen during PM off-peak, followed by PM peak, and midday, whereas AM off-peak hours are the least busy. In our data, an average driver works 2.1 days per week and 6.35 hours per day.

In this paper, we analyze drivers' behavior at both the shift and day levels. We control for the day of the week to account for demand and supply variation. In our data, 49.46% of all completed trips occurred between Tuesday and Thursday, potentially confirming the popularity of the service among city commuters. Monday and Friday trips account for 30.91% of all trips, whereas weekend trips account for 19.62%. Although drivers are allowed to flexibly decide their own work schedules, they often stick to their regular times. For example, 30.41% of drivers never worked on weekends, and 91.07% of drivers' working days did not overlap with midnight (e.g., they did not work overnight).

3.3. Earnings and Incentives

Drivers receive a shift-specific hourly rate for the duration they are active on the platform. They are considered active when they log on to the driver application on

Figure 1. (Color online) Breakdown of Shifts for Each Operating Day

their mobile device and report to their designated start location. This compensation scheme can be considered as a guaranteed payment, in contrast to a commission-based contract that compensates drivers for each completed trip, which is commonly used by several platforms. It is possible under this scheme that drivers could be paid even if there are no ride requests for the entire hour.[2] Similar schemes are used by other gig platforms such as DoorDash, GoPuff, and HourlyBee.

The guaranteed hourly offer comprises two components: a base rate and a promotional rate. These two rates vary over time (shifts and days of week) and across different drivers. The base rate for each driver is decided when the driver joins the platform for the first time. For the same driver, the base rate may be different across shifts and across days of the week, but typically remains the same across weeks. In addition to the base rate, drivers are frequently offered promotional incentives. Rate-based promotions provide a multiplicative bonus to the hourly base rate during specific times (e.g., during 2× shifts, drivers earn twice the base rate); 32.71% of shifts in the data include rate-based promotions and the average promotion rate is an additional 50.36% of the base rate or approximately 1.5×.

At the time of our data, incentives were decided as follows: First, the platform sets a number of promotional rates as benchmarks. Then, an algorithm uses these rates to assign the final rate for each driver based on recent work history and vehicle type. Both the base and promotional rates are specific to each driver. The platform then sends text messages to drivers every evening to communicate the rates for the following day. This suggests that drivers are likely to plan their work schedule ahead of time and there is no internal competition for better rates among drivers. Occasionally, drivers may receive real-time adjustments to their rates but will never experience lower rates than initially informed. All rates are prorated to the actual amount of time worked in a given shift. Earnings are cumulative until the end of the week when drivers have the option to transfer their earnings to their bank account.

3.4. Drivers and Vehicle Types

Drivers are identified by a unique ID. For each shift, we observe the decision to work (i.e., to become active) for every driver registered in the system. For drivers who started working after the first day of our data set, we record both their first day joining the platform and their first workday to control for their experience with the platform. Similarly, we observe the last day of being registered with the platform for some drivers if they left within the duration of our data. This allows us to control for drivers' experience, tenure, and span of their service for the focal platform.

For the analysis conducted in this paper, we only consider the drivers who own a single vehicle (89.9% of all drivers). There are six types of vehicles: a three-passenger sedan, a small three-passenger sports utility vehicle (SUV), a medium four-passenger SUV, a large five-passenger SUV, a five-passenger van, and a six-passenger van. We exclude van drivers from our analysis as the majority of them lease their vehicle from the company rather than own their vehicle or lease it from an external third party, leaving us with 86.3% of the original pool of drivers. For our main analysis, we present the results for two types of vehicles: sedan and large SUV, which are 33.2% of the pool. We make an assumption that drivers of different vehicle types may have fundamentally different utilities and preferences. Sedan vehicles are generally less expensive to maintain than SUVs, whereas SUV drivers may have a different set of outside opportunities (e.g., qualified for both regular and XL services). From our data, we observe that SUV drivers typically work more frequently and for longer hours relative to sedan drivers. We obtain similar qualitative results for other vehicle types; but omit them for conciseness.

3.5. Supplementary Data: TLC Trip Records

We incorporate trip records for other similar services in the same region to capture the real-time market conditions. Information about taxi and for-hire vehicle (FHV) trips in New York City has been collected by the Taxi and Limousine Commission (TLC) and publicly released since 2009.[3] In particular, we analyze 101,487,565 yellow taxi trips and 129,868,077 FHV trips operated by four major service providers (including our focal platform) in the city between October 2016 and September 2017 (i.e., the duration of our data). Taxi trip records include date, time, and location (at the neighborhood level) of every pick-up and drop-off, itemized fares, and driver-reported passenger counts. FHV trip records prior to July 2017 consist of date, time, and location of each pick-up and the dispatching base associated with a ride-hailing platform. Starting from July 2017, we also observe date, time, and location of each drop-off by FHV drivers. In Section 4.1, we discuss the metrics that we construct to control for market conditions and competition intensity.

4. Empirical Approach

To test the hypotheses developed in Section 2, we estimate the impact of financial incentives, income and time targets, and other covariates on two labor decisions: (i) *whether to work or not* and (ii) *work duration*. We assume that drivers make both decisions at the beginning of each shift or day. We conduct our analyses at two levels, within-day (shift level) and across-days (day level), as well as for each vehicle type separately. This allows us to understand how variations within the same day or across days affect drivers' decisions and to capture vehicle type-specific heterogeneity. Drivers operating different vehicle types may have different preferences, costs, and utility functions, and thus make their labor decisions differently. In this section, we first introduce our econometric model and key covariates, then provide details of our estimation method, and finally discuss the empirical challenges and our strategies to address them.

4.1. Empirical Model and Estimation Details

As discussed, our data set provides a unique advantage as we observe the financial incentives offered to every driver for every shift as long as they already joined the platform and have not yet terminated their drivership. This allows us to study two stages of labor decisions and control for potential sample selection bias (see Section 4.2.1 for further discussion). Our approach therefore adapts the two-stage Heckman estimation method (Heckman 1979) to first estimate the decision to work across all drivers using a probit regression, and then estimate the work duration for drivers who chose to work for any given shift or day using an ordinary least squares (OLS) regression.

4.1.1. Outcome Variables. The decision of the first stage is captured by the binary variable $Drive_{i,t}$. Specifically, $Drive_{i,t} = 1$ if driver i works during shift (or day) t and $Drive_{i,t} = 0$ otherwise. In the second stage, conditional on working during shift (or day) t, $Hours_{i,t}$ represents work duration in hours for driver i during t. Given the long tails in $Hours_{i,t}$, we apply a Box-Cox transformation conditional on the covariates to normalize its distribution and homogenize its variance. Our results are robust under other types of transformation (e.g., logarithm, square root) and also without a transformation. We exclude outliers defined as drivers whose work duration during a given shift or day exceeds the 1.5 interquartile ranges (IQRs) or less than five minutes. We also exclude public holidays from our analysis.

4.1.2. Key Covariates. We focus our analysis on three key drivers of labor decisions. (i) Financial incentives. We use the hourly offer rate (i.e., the sum of hourly base rate and promotions, if available), denoted as $w_{i,t}$ for driver i during shift (or day) t, for the first stage. Similarly, conditional on working, the second stage's financial incentives are taken from the hourly earnings rate (i.e., the sum of hourly base rate and promotions, if available), denoted as $\tilde{w}_{i,t}$. (ii) Income targets. As we do not directly observe drivers' income targets, we use cumulative earnings since the beginning of the day (week) until the focal decision point as a proxy for a daily (weekly) income target. We refer to this covariate as *income so far* or *ISF*. The rationale behind this proxy is that, as drivers start accumulating earnings, the higher *ISF*, the closer they are to their privately known targets. The same proxy is used in the literature (e.g., Crawford and Meng 2011, Thakral and Tô 2021). (iii) Time targets. Similarly, we use cumulative work hours since the beginning of the day (week) until the focal decision point as a proxy for a daily (weekly) time target. We refer to this covariate as *hours so far* or *HSF*. Given our observation that over 90% of the data do not include overnight work, we assume that daily targets and progress are reset at midnight (e.g., the driver starts working toward a new target for the new day). Similarly, as the majority of work occurred during weekdays, we assume that weekly targets are reset at the end of every Sunday. Our results are robust to different constructs of targets and flexible frequency of target reset.

4.1.3. Two-Stage Estimation. Let $w_{i,t}, \tilde{w}_{i,t}, ISF_{i,t}$, and $HSF_{i,t}$ be hourly offer, hourly earnings rate, cumulative income, and cumulative work hours of driver i at the beginning of time t, respectively. The variables $\boldsymbol{X_{i,t}}$ and $\boldsymbol{Z_{i,t}}$ are other relevant covariates that affect the decision to work and work duration, respectively. We model the two stages of labor decisions, $Drive_{i,t}$ and $Hours_{i,t}$, of driver i at time t as follows:

$$Hour_{i,t} = \begin{cases} Hour^*_{i,t} & \text{if} \quad Drive_{i,t} = 1 \\ \text{unobserved} & \text{otherwise}, \end{cases} \tag{1}$$

$$Drive_{i,t} = \begin{cases} 1 & \text{if} \quad Drive^*_{i,t} > 0 \\ 0 & \text{otherwise}, \end{cases} \tag{2}$$

$$Drive^*_{i,t} = \alpha_{0,i} + \alpha_w w_{i,t} + \alpha_{ISF} ISF_{i,t} + \alpha_{HSF} HSF_{i,t} + \boldsymbol{\alpha X_{i,t}} + v_{i,t}, \tag{3}$$

$$Hour^*_{i,t} = \beta_{0,i} + \beta_{\tilde{w}} \tilde{w}_{i,t} + \beta_{ISF} ISF_{i,t} + \beta_{HSF} HSF_{i,t} + \boldsymbol{\beta Z_{i,t}} + u_{i,t} \tag{4}$$

$$\begin{bmatrix} \sigma_v^2 \\ \sigma_u^2 \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho\sigma_u \\ \rho\sigma_u & \sigma_u^2 \end{bmatrix} \right). \tag{5}$$

The two stages that we estimate are given by:

$$P(Drive_{i,t} = 1 \mid \boldsymbol{X_{i,t}}) = \Phi(\alpha_{0,i} + \alpha_w w_{i,t} + \alpha_{ISF} ISF_{i,t} + \alpha_{HSF} HSF_{i,t} + \boldsymbol{\alpha X_{i,t}}), \tag{6}$$

$$f(Hour_{i,t}) = \beta_{0,i} + \beta_{\tilde{w}} \tilde{w}_{i,t} + \beta_{ISF} ISF_{i,t} + \beta_{HSF} HSF_{i,t} + \boldsymbol{\beta Z_{i,t}} + \theta\lambda_{i,t} + u_{i,t}, \tag{7}$$

where $\Phi(\cdot)$ is the normal cumulative distribution function (c.d.f.) and $\lambda_{i,t}$ is the inverse Mills ratio (IMR) calculated from the predicted probability in Equation (6) (choice equation). Thus, we essentially estimate a probit model for the work decision in Equation (6) and compute the IMR for each observation. We then fit an OLS model of the (transformed) work duration conditional on all covariates and the IMR (Equation (7)), while controlling for the drivers who worked (level equation). The estimated coefficient $\theta = \rho\sigma_u$ will potentially confirm the existence of a sample selection bias. We next discuss in detail the estimation methodology for each stage.

4.1.3.1. Choice: Control Function Probit. The first stage is based on a probit model of labor decisions, $Drive_{i,t}$.

We address a potential endogeneity related to financial incentives and past work decisions by taking an instrumental variable (IV) approach (see Section 4.2.2). A commonly used two-stage least squares (2SLS) can provide inconsistent estimates for a probit model as certain properties of the expectation and linear projection operators do not carry over to nonlinear models (Newey 1987). Instead, we implement the control function method to account for endogeneity for our nonlinear probability model (Imbens and Wooldridge 2007, Wooldridge 2015). The first step is identical to the first step of 2SLS, that is, we estimate an OLS regression of the endogenous variable ($w_{i,t}$) on exogenous covariates and instrumental variables. We can then keep the endogenous variable in the model and include the residuals from the previous regression as an additional regressor. The intuition behind this method relies on using the instrument to split the unmeasured confounders into two parts, one that is correlated with the endogenous regressor and one that is not. We correct for the standard errors using the standard deviation of the residuals following Imbens and Wooldridge (2007).

We also allow for drivers and time fixed effects throughout our estimation. Adding fixed effects to the nonlinear choice equation is known to generate the incidental parameters problem. More precisely, the usual asymptotic properties of the maximum likelihood estimator are not guaranteed, thus leading to a biased and inconsistent estimator (Greene 2004). Fortunately, recent developments in bias correction, such as the jackknife estimation method (see Hahn and Newey 2004, Dhaene and Jochmans 2015 for more details on this method), allow us to obtain asymptotically unbiased estimates and alleviate the incidental parameters problem. The final step for this stage is to compute the IMR for each observation using the fitted probability.

4.1.3.2. Level: Fixed Effects 2SLS. The second stage aims to estimate the work duration, $Hour_{i,t}$, conditional on the driver working during the focal time period. The hourly earnings rate, $\tilde{w}_{i,t}$, is likely to be endogenous. Incorporating the IV approach to the level equation is straightforward, as we can simply perform a 2SLS regression in which we first obtain the predicted value of $\tilde{w}_{i,t}$ based on exogenous covariates and the IVs. We transform the observed work duration using a Box-Cox approach conditional on all covariates to alleviate heteroskedasticity. Finally, as we include the IMR as one of the regressors in the second stage, we bootstrap the standard errors by repeating our analysis on resampled datasets.

4.1.3.3. Other Covariates. To capture drivers' heterogeneity, we first include a driver-specific intercept in both stages, even if we already perform separate analyses for drivers with different vehicle types. We also include other time-varying driver-specific covariates that could reflect their work habits. Short-term habits are captured by historical work duration on the same day and shift of the previous week and the total hours worked during the previous week. Long-term habits are captured by the driver's experience (i.e., whether they are new to the platform and their tenure) and also through drivers' fixed effects. Month and day-of-week fixed effects are also included to capture seasonal trends. The sets of regressors in our main model are:

- **Choice:** Hourly offer (w), cumulative earnings (*ISF*), cumulative work hours (*HSF*), number of hours worked last week, new driver indicator, humidity, apparent temperature, precipitation probability, and number of other ride-hailing trips in the previous shift or day (in thousands).
- **Level:** Hourly earning rate ($\tilde{w}$), cumulative earnings (*ISF*), cumulative work hours (*HSF*), number of hours worked on the same shift of last week, humidity, apparent temperature, precipitation probability, and number of other ride-hailing trips during the same shift or day (in thousands).

4.2. Empirical Challenges and Strategies

4.2.1. Sample Selection Bias. Previous studies such as Camerer et al. (1997) and Sheldon (2016) investigated the relationship between the number of work hours and the hourly wage conditional on drivers who worked on a given day. This would not be a concern if drivers randomly decide whether to work or not. In reality, however, it is more plausible that they make such decisions based on factors that are not observed by the researcher. In other words, the selection of drivers who choose to work at a given time is not random. Consequently, this approach may yield a biased estimate of the sensitivity to incentives (i.e., income elasticity). Fortunately, the comprehensiveness of our data offers an opportunity to address this challenge. Because we observe incentives for all drivers on every shift regardless of their work decisions, we can directly estimate the selection problem. As presented in Section 4.1.3, we employ a modified two-stage Heckman estimation method for our analysis.

Whereas the Heckman-type selection model has been widely used in several applications, it has also been criticized on its potential pitfalls, particularly the weak nonlinearity of the IMR and the multicollinearity of regressors in both stages (Puhani 2000). To address these concerns, we carefully choose the sets of regressors for both stages ($\boldsymbol{X}_{i,t}$ and $\boldsymbol{Z}_{i,t}$) to be different (as shown in Section 4.1.3) and we check for collinearity by regressing the IMR on the regressors of the second stage. On average, the standard deviation of the errors is 44.52% less than the standard deviation of the IMR, which suggests a substantial difference. We also consider an alternative approach suggested by Puhani (2000): estimating a subsample OLS or a two-part model. In the two-part model, a binary choice model is estimated for the probability of observing a positive-versus-zero outcome (e.g., the number of work hours). This is essentially the same as the first stage of our main approach. Conditional on a positive outcome (e.g., drivers who worked during a particular shift or day), a separate OLS regression model is estimated for the work duration (Cragg 1971, Madden 2008, Farewell et al. 2017). This is the same as the second stage of our main approach, excluding the IMR. We report the estimates from both the two-part model and our main approach in Section 5. Finally, as a robustness check, we consider Dahl's approach by using a basis spline to approximate the choice probability (Dahl 2002). For more details on the approach, we refer the reader to Bourguignon et al. (2007), which provides Monte Carlo comparisons across different selection models and to Bray et al. (2019), which implements this correction to model proximity-based supplier selection. In our context, the choice for each driver is binary. Our results remain consistent and are presented in online Appendix C.1.

4.2.2. Endogeneity. As discussed in Section 2.1, the standard income effect suggests that financial incentives encourage workers by increasing their likelihood of working or work duration. Nevertheless, quantifying the effect of incentives by regressing the labor decision on financial incentives can lead to misleading results. In our data set, we observe that a smaller fraction of drivers who received an hourly offer of $65 decided to work relative to those who received $45 per hour. One possible implication is that financial incentives are not effective in convincing some drivers. Alternatively, these appealing promotions might have been strategically offered to engage inactive drivers. Consequently, regressing work decisions on financial incentives can lead to an omitted variable bias as we do not observe the actual algorithm behind these incentives. Overlooking this issue may yield to a bias estimate of the effect of financial incentives. A common solution is to use instrumental variables (IVs) that are highly correlated with financial incentives but affect the work decision only through the incentives (Levinsohn and Petrin 2003).

4.2.2.1. Instrumental Variables. The main endogenous variables in our data are the hourly financial incentives, $w_{i,t}$, and the hourly earnings, $\tilde{w}_{i,j}$. Our ideal instrument should be highly correlated with each endogenous variable and affect the dependent variable (the decision to drive or the work hours) only through the endogenous variable. In other words, we are looking for instruments that are not correlated with the unobserved variables in the error terms. Our industry partner confirmed that the financial incentives were endogenously determined with respect to (predicted) supply decisions. Specifically, the firm sets the incentives based on past work history, level of inactivity, and vehicle type. Different teams are in charge of determining the offers for different vehicle types. This insight motivated us to focus on instruments that categorize drivers based on these three factors.

Our instrument is based on the notion of coworkers. For each driver who is available to work at a particular time (i.e., has not terminated his or her partnership with the platform), we define the driver's coworkers as the drivers who meet the following conditions: (i) available to work at the same time, (ii) drive a different vehicle type, and (iii) have made the same work decision in the past (i.e., worked in the same shift in the previous week or previous month). Work decisions are binary such as working or not. Assuming that random shocks, $v_{i,t}$ and $u_{i,t}$, are not correlated across drivers, we propose to use the average hourly offers received by coworkers for the focal period as an IV. This IV satisfies the relevance condition: because both the focal driver and the driver's coworkers made the same work decision in the past, their incentives should be highly correlated as the firm would adjust the incentives for both groups in the same direction. From the first stage of our IV estimation, the estimate for the instrument is consistently significant and F-statistics for all models are higher than the conventional threshold of 10. This IV also satisfies the exclusion restriction: current incentives for coworkers should not directly influence the focal driver's work decision because (i) the offers for different vehicle types are decided independently by different teams within the company, (ii) the focal driver does not have access to coworkers' incentives information, and (iii) it is unlikely that drivers compare the offers across different vehicle types.

To test the robustness of our results, we consider two alternative instruments. First, instead of matching drivers based on their decision to work at a specific time in the past, we now match drivers based on their decision not to work: the level of past inactivity. For every day in our data, we categorize drivers into four groups based on each quartile of the number of consecutive days they have not been working. We refer to the drivers of a different vehicle type who belong to the same group as *coskippers*. Finally, we also consider the instrument used in previous literature (e.g., Sheldon 2016), the average hourly offer rate received by all other drivers during the same shift on the same day as an instrument for the offer rate. We obtain consistent insights under all three specifications. Further details are deferred to Online Appendix C.2.

4.2.3. Multicollinearity. A potential concern of including both *HSF* and *ISF* in the same specification is the multicollinearity issue. Correlations between *HSF* and *ISF* in our data range between 0.667 and 0.928, depending on the time of the day and the vehicle type. This issue does not significantly affect our results because of three reasons. First, despite a positive correlation, *HSF* and *ISF* are not a direct transformation of each other, hence there is no perfect correlation. Intuitively, *HSF* increases linearly with time as it denotes the exact amount of time the driver has been working, whereas *ISF* evolves dynamically as it depends on time-varying financial incentives. Second, multicollinearity generally makes causal inference difficult as the variance of each estimate would be inflated, leading to statistical insignificance, but the estimate itself would be unbiased. Our main results (see Section 5) show that this is not the case for us as both coefficients for *HSF* and *ISF* are statistically significant in most cases. Third, potential problems from high collinearity can be largely offset with sufficient power (Mason and Perreault 1991). Our data set consists of a large enough number of observations to provide sufficient statistical power even when we separately estimate our model by vehicle type, day of the week, and shift of the day. Finally, we consider several alternative approaches to alleviate the multicollinearity concerns, including performing localized regressions by controlling for drivers with similar *HSF* or *ISF*, considering models with only one of the two proxies (*HSF* or *ISF*), and converting one of the two variables to be categorical. Our insights remain qualitatively consistent in all cases. Further details and discussion are deferred to Online Appendix C.3. We also note that the correlations between *HSF*/*ISF* and the hourly offer rates are weak and not statistically significant.

4.2.4. Competition with Other Ride-Hailing Platforms.

One of the key features of the gig economy is the flexibility that gig workers have in choosing their work schedule as well as the platform to work for. During the timeframe of our data set, there were four major ride-hailing companies operating in NYC. All ride-hailing drivers require a TLC license plate to work in the five city boroughs. Drivers on our focal platform are therefore eligible to work and could have worked for other companies and made these choices during the same time as our data. Capturing the outside options of each driver is thus crucial in understanding their labor decisions. The main challenge is that we do not observe when drivers from our focal platform could have worked for other companies nor the information about incentives outside our focal platform. In our main specification, we include two covariates that can shed some light on the current market conditions for ride-hailing services. First, we capture the recent volume of rides operated by the ride-hailing competitors using the number of trips from the TLC trip records data. In the choice equation, we include the number of trips on competing platforms initiated in the previous period, $NumFHV_{t-1}$, to reflect the market condition observed by the drivers in our platform at the time of decision t. Second, we capture the current volume of competing services in the level equation by using the number of trips initiated in the same period, $NumFHV_t$.

We create two metrics to capture competition effects by leveraging additional information on drop-off time and location of all FHV drivers as well as the trip distance and duration of taxi drivers (which is only available starting from July 2017). First, to capture the traffic and congestion conditions, we compute the speed (in miles per hour) for each taxi trip by dividing the trip distance by the trip duration. We then compute the average speed for trips initiated in each neighborhood at each time period. To match with a shift (or day) in our data, we average across all neighborhoods and time periods within the shift (or day). We then include the average speed, $Speed_t$, in both stages. Second, to reflect potential real-time adjustments to financial incentives (e.g., surge pricing) on competing platforms, we compare the imbalance between supply and demand in each neighborhood at each time period. We assume that drivers who recently dropped off passengers in the neighborhood reflect the number of potential supply of drivers in that neighborhood. In the same vein, if we observe a larger number of trips picking up passengers from a specific neighborhood, we can infer that this neighborhood has high demand (compared with supply), and hence would likely trigger surge prices on the competitors' platforms. We define the binary variable $Surge_{l,t}$ as whether the number of trips leaving location l is at least 1.5 times greater than the number of trips entering the same location at time t. In other words, surge pricing is likely to be activated when there are at least 50% more ride requests than the number of available drivers in the neighborhood. Using different thresholds yields qualitatively similar insights. We then compute the number of neighborhoods in the city with $Surge_{l,t} = 1$ for each time t. Aggregating across hours to a shift level, we obtain $AggSurge_s = \sum_{t \in \text{Shift}_s} (\sum_{l \in \mathcal{L}} Surge_{l,t}) / |\mathcal{L}|$ as our metric for potential real-time appealing opportunities for the drivers to work for the competing platforms during shift s, where $\mathcal{L}$ is a set of neighborhoods in NYC. Our insights remain valid with the inclusion of these metrics. Details and discussion of the results are presented in Online Appendix D.

5. Empirical Results

We first present our analysis at the shift level, understanding the impact of financial incentives, income, and time targets on within-day labor decisions of SUV and sedan drivers. We then perform the analysis at the day level, to study across-day labor decisions from Tuesday to Sunday. We discuss the insights from both analyses and test the hypotheses developed in Section 2. Finally, we discuss a number of robustness tests that help validate our findings.

5.1. Within-Day Analysis

We examine drivers' labor decisions at the beginning of each of the company-specified shifts as introduced in Section 3.2. As 91% of drivers' working days observed in our data do not overlap with midnight and 73% of workday happened between 7 a.m. and midnight, we assume that the first shift of the day is AM peak (starting at 7 a.m.) and the last shift of the day is late night (ending at midnight). Our analysis focuses on four shifts (midday to late night) to investigate how labor decisions are influenced by financial incentives (*offer*) as well as by cumulative earnings (*ISF*) and work hours (*HSF*) since the beginning of the day. We assume that daily income and time targets, proxied by *ISF* and *HSF*, are reset every day after midnight.

For each shift, we first estimate the choice equation (Equation (6)) in which the outcome variable is a binary decision of whether to work for the focal shift using a control function approach. We then estimate the level equation (Equation (7)) that concerns the work duration for the shift, conditional on the decision to work. Figure 2 summarizes the signs and statistical significance of the key estimates (hourly offer/earnings, *ISF*, and *HSF*) for each vehicle type and each shift. Each cell in the main three columns reports the sign of the effect (+ or –) and its statistical significance at $p = 0.05$ as follows: solid background with a bolded + indicates a significant positive estimate, striped background with a bolded – indicates a significant negative estimate, and white background with italicized sign corresponds to a nonsignificant directional effect. In addition, we provide the mean work probability, *F*-statistics from the first stage of each IV estimation, mean work duration conditional on working, adjusted total R^2, and number of observations alongside the estimates. We report detailed results as well as a comparison of our main model with two other specifications for the second stage: a baseline OLS and a 2SLS without correction for sample selection bias (two-part model) at a shift level in Online Appendix B.

We observe that the estimates for drivers of both vehicle types are similar across most shifts. Hourly offers have a consistent positive impact on both choice and level decisions. This result is consistent with the standard income effect that predicts a positive income elasticity and confirms our first hypothesis, namely that financial incentives encourage the decision to work and boost the work duration. However, we also observe evidence of behavioral factors of labor supply with regard to cumulative earnings and work hours. The impact of *ISF* on both stages is significantly negative, suggesting that drivers become less likely to work and will work for shorter durations when they have earned higher cumulative income since the beginning of the work day. This phenomenon reflects an income-targeting behavior among drivers and provides support that labor decisions are negatively influenced by an income-targeting behavior, hence supporting our second hypothesis. Lastly, we observe a somewhat surprising effect from *HSF* on both stages. Specifically, drivers who have previously worked for a longer duration since the beginning of the day are more likely to work in a new shift and for a longer duration. We refer to this phenomenon as *inertia*. Our third hypothesis is hence rejected in the sense that, when controlling for the key covariates, drivers do not exhibit a time-targeting behavior or an aversion to working too many hours.

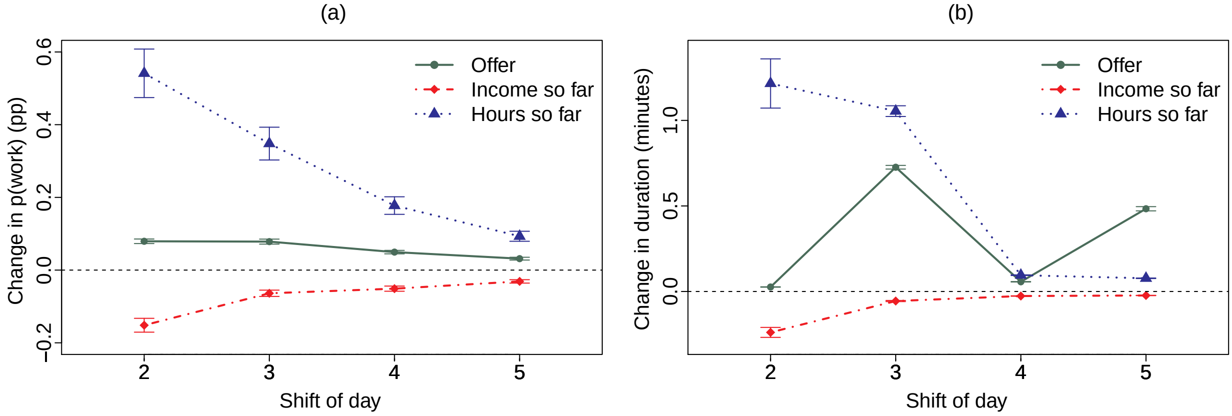
As our three key variables have different units, it is not straightforward to compare the magnitude of their effects. Nevertheless, we can compare how the probability of working and the work duration are affected by a 1% increase in each of the variables for an average driver. Figure 3(a) and (b) illustrate the change in probability of

Figure 2. (Color online) Signs and Statistical Significance for Estimates of Two-Stage Models of Drivers' Shift-Level Decisions

	Choice (Work or not)						**Level** (How long)						
SUV	Mean	IV-F	Offer	ISF	HSF	N	Mean	IV-F	Earn	ISF	HSF	R^2	N
Midday	0.343	372.9	**+**	**-**	**+**	124,769	4.987	180.2	*+*	**-**	**+**	0.552	45,329
PM-peak	0.277	345.1	**+**	**-**	**+**	131,910	2.421	58.5	**+**	**-**	**+**	0.244	39,592
PM-off-peak	0.182	320.6	**+**	**-**	**+**	130,651	0.731	50.5	**+**	**-**	**+**	0.281	26,699
Late night	0.117	379.0	**+**	**-**	**+**	125,382	1.996	39.91	**+**	*-*	**+**	0.296	17,137
Sedan	Mean	IV-F	Offer	ISF	HSF	N	Mean	IV-F	Earn	ISF	HSF	R^2	N
Midday	0.137	224.9	**+**	**-**	**+**	113,444	4.186	78.0	*+*	**-**	**+**	0.580	20,297
PM-peak	0.123	254.7	**+**	**-**	**+**	117,152	2.327	32.6	**+**	**-**	**+**	0.273	19,613
PM-off-peak	0.099	298.8	**+**	**-**	**+**	124,611	0.803	29.9	**+**	**-**	**+**	0.252	17,025
Late night	0.071	299.5	*+*	**-**	**+**	124,280	2.167	32.9	**+**	**-**	**+**	0.304	15,623

Notes. Solid background with bolded +: significantly positive; striped with bolded −: significantly negative; white with italicized sign: nonsignificant. All at $p = 0.05$.

Figure 3. (Color online) Change in Outcome for an Average SUV Driver When Each Variable Increases by 1%: (a) Change in Probability of Working in Percentage Points; (b) Change in Work Duration in Minutes



working (in percentage points) and the change in work duration (in minutes) from midday to late night for an average SUV driver, respectively. During earlier shifts in the day, the marginal effect of *HSF* dominates that of the hourly offer and *ISF*. We also observe that the behavioral effects (e.g., income targeting and inertia) are weaker later in the day. The detailed effect sizes for both SUV and sedan drivers are reported in Online Appendix B.

Putting these together, we conclude that drivers exhibit positive income elasticity as predicted by the standard income effect but are also influenced by behavioral motives such as income targeting and inertia.

5.2. Across-Day Analysis

Here, we consider the labor decisions that drivers make at the beginning of each day, whether to work for the day and, if so, for how long. We assume that the week starts on Monday so the income target *ISF*, the time target *HSF*, and their progress are reset at the end of Sunday. In this analysis, *ISF* and *HSF* are therefore considered as proxies for the weekly income and time targets. The covariates in both stages of the estimation are nearly identical to the ones used in Section 5.1, except that we replace the past work duration on the same shift of the previous week by the past work duration on the same day of the previous week. Figure 4 displays the estimates from our model for both vehicle types.

At a day level, we draw considerably different conclusions from our shift-level analysis. Whereas the positive impact of *HSF* on a decision to work remains consistent, the impact of hourly offer and *ISF* appear to vary across days of the week. Prior to the weekend, both hourly offer and *ISF* positively encourage drivers to work. The latter effect might suggest that drivers perceive high cumulative earnings early in the week as an indicator of high demand and form an optimistic outlook on future market conditions. However, both effects become negative for Saturday and Sunday, resembling less effectiveness of financial incentive and weaker income-targeting behavior. The results for the level equation shed another interesting insight. We do not find significant effects from the three main drivers in most cases, except a consistent inertia observed among sedan drivers. Note that the estimates of the IMR are significant across all cases, suggesting that there is indeed a sample selection bias in the daily work decision. One potential explanation is that, whereas gig economy workers make strategic decisions of whether to work on a daily basis, they do not seem to decide the work duration for the entire day ahead of time. Instead, they are likely to make such a decision at the shift (or hour) level as observed in our shift-level analysis.

5.3. Discussion

Our results offer a refined explanation of how gig economy workers make labor decisions and, in part, reconcile the debate between neoclassical and behavioral theories of labor supply. Table 1 summarizes our hypotheses and results. We find that, as predicted by the standard income effect, drivers respond positively to financial incentives. Although we do not observe the strong negative income elasticity from the literature (such as Camerer et al. 1997), we find empirical evidence of an income-targeting behavior among drivers, suggesting that their labor decisions are influenced by recent earnings or income goals. Several gig economy platforms provide in-app features such as a real-time progress dashboard, making it simple for workers to track their progress and recent earnings and work history. In other words, information surrounding past earnings and work activities has become much more salient relative to traditional settings. By separating cumulative income from financial incentives, we show that the negative impact of income targeting stems from

Figure 4. (Color online) Signs and Statistical Significance for Estimates of Two-Stage Models of Drivers' Day-Level Decisions

	Choice (Work or not)						Level (How long)						
SUV	Mean	IV-F	Offer	ISF	HSF	N	Mean	IV-F	Earn	ISF	HSF	R^2	N
Tuesday	0.409	43.6	*+*	*+*	**+**	28,883	8.696	18.3	-	-	*+*	0.422	9,482
Wednesday	0.418	55.9	*+*	**+**	**+**	21,965	8.964	26.2	-	-	*+*	0.422	10,120
Thursday	0.426	73.4	**+**	**+**	**+**	29,233	9.053	34.6	-	-	**+**	0.412	9,894
Friday	0.412	74.0	**+**	*+*	**+**	20,294	8.915	33.7	*+*	-	**+**	0.436	9,283
Saturday	0.203	98.1	**-**	-	**+**	15,788	8.435	19.1	-	*+*	-	0.398	4,372
Sunday	0.162	82.2	**-**	**-**	**+**	13,025	7.927	15.1	**+**	*+*	-	0.390	3,240
Sedan	Mean	IV-F	Offer	ISF	HSF	N	Mean	IV-F	Earn	ISF	HSF	R^2	N
Tuesday	0.169	31.1	**+**	*+*	**+**	21,283	7.687	7.3	**-**	-	**+**	0.564	4,681
Wednesday	0.182	37.3	**+**	**+**	**+**	23,280	7.680	9.8	*+*	**-**	**+**	0.567	5,278
Thursday	0.179	47.5	**+**	**+**	**+**	19,982	7.724	11.6	-	-	**+**	0.542	5,081
Friday	0.171	46.7	**+**	**+**	**+**	18,418	7.568	11.2	-	-	**+**	0.533	4,666
Saturday	0.148	53.3	*+*	-	**+**	15,762	8.022	11.7	-	-	**+**	0.514	3,817
Sunday	0.129	45.5	-	**-**	**+**	12,602	7.708	11.4	**+**	**-**	**+**	0.560	3,065

Notes. Solid background with bolded text: significantly positive; striped with bolded text: significantly negative; white with italicized text: non-significant. All at $p = 0.05$.

cumulative income rather than the hourly wage. Thakral and Tô (2021) similarly demonstrate the existence of income targeting among taxi drivers and identify the recently earned cumulative income as a key factor in the decision to quit.

In addition, we find that workers who have previously worked for a longer duration are more likely to start a new shift and work for a longer period of time compared with the drivers who have recently worked less, controlling for all other covariates. We refer to this finding as inertia to reflect the tendency of workers with longer recent work hours to continue working and stay active for longer than their counterparts. Our result on inertia is in contrast with findings from Crawford and Meng (2011) and Farber (2015) that taxi drivers exhibit a time-targeting behavior. This difference could be driven by the unique flexibility of gig work. Inertia could represent drivers' strategic behavior related to consistency and learning. In one of our additional analyses to identify underlying mechanisms that drive inertia and long work hours, we include the interaction terms between drivers' work experience and each of the key variables. We find that both income targeting and inertia are significant among drivers with less experience (e.g., fewer than 80 working days). The directions of these effects still hold for drivers with more experience. A similar impact of experience is documented by Sheldon (2016). We also find that inertia is more prevalent on days when the financial offers on the focal platform have a high variance or when the average competition intensity is high but the variance is low. Detailed discussion and results of these analyses are reported in Online Appendix E. Furthermore, multiple psychological phenomena could potentially explain the existence of inertia, such as reduced fatigue from voluntarily scheduled work (Beckers et al. 2008) and work addiction driven by stochastic and frequent rewards (DeVoe et al. 2010, Corgnet et al. 2020). We also believe that workers' different behaviors toward

Table 1. Summary of Hypotheses and Results

		Shift level		Day level	
Hypothesis	Statement	SUV	Sedan	SUV	Sedan
1a	Higher wage increases P(work)	✓	✓	✓ → ✗	✓ → ✗
1b	Higher wage increases work hours	✓	✓	✗	✗
2a	Higher income so far decreases P(work)	✓	✓	✗ → ✓	✗ → ✓
2a	Higher income so far shortens work duration	✓	✓	✗	✗
3a	Longer work hours so far decreases P(work)	✗	✗	✗	✗
3b	Longer work hours so far shortens work duration	✗	✗	✗	✗

Note. P(work): likelihood of working; ✓: fail to reject; ✗: reject; →: result differs later on in the day or week.

time versus money could be explained by how people perceive the value of time and money differently. Psychological research has found that mental accounting for time does not work in the same manner as mental accounting for money (Leclerc et al. 1995, Soman 2001). See Online Appendix F for further discussion on this topic. It could also be possible that inertia arises from error correlation within each driver over time for our shift-level analysis. However, with our shift-specific estimation and assumption that errors for the same shifts (e.g., midday) across different days are independent, potentially unobserved driver-specific shocks such as illness are likely to be short-lived (e.g., spanning adjacent shifts within the same day) and would be unlikely to affect our estimates (e.g., same shifts across days). For our day-level analysis, we assume that errors for the same days of week across weeks are independent. Therefore, the day-level error correlation should be even smaller and less likely to affect our results.

Lastly, we find that gig workers make a decision to work at both shift and day levels, whereas the work duration appears to be decided at a more granular time unit such as a shift or even an hour. The latter potentially highlights the unique flexibility of gig jobs that provide workers with full control of their real-time work schedule. Our results remain valid under a number of robustness checks, such as the following: allowing for nonlinear targeting effects, relaxing our assumption on frequency of target adjustment and definition of shifts, considering instrumental variables for *ISF* and *HSF*, performing alternative sample selection correction, and modeling stopping probabilities via localized hazard regressions and mixed-effects survival analysis (see Online Appendix C). In summary, with a better understanding of how gig workers make labor decisions, companies can design more effective incentives and personalize them based on individual workers' behaviors.

6. Managerial Implications: Optimal Incentive Allocation

In this section, we illustrate how gig economy firms can use our insights on workers' behavior to enhance their operations. We first investigate the benefit of improved incentive allocation based on two perspectives: (i) increasing service capacity while keeping a fixed budget and (ii) maintaining the same service capacity at a lower cost. We then further highlight the potential pitfalls of ignoring behavioral factors and quantify the resulting capacity loss. In Online Appendix G, we conduct a policy analysis to demonstrate how our insights can help policymakers evaluate the impact of a regulation.

6.1. Targeted Incentives

Our main results suggest that workers are influenced by their behavioral motives and that the impact of incentives on the number of active workers may be nonlinear. Targeting specific workers with different incentives can be beneficial. We examine how the platform can improve its operational performance by offering personalized incentives based on workers' attributes. As a benchmark, we compute the platform's budget for promotions based on the actual allocation of incentives. We then reallocate the promotion budget more efficiently by considering the following two perspectives: (i) increasing the service capacity (i.e., staffing more workers) using the same budget, and (ii) maintaining the same service capacity at a lower cost. Our proposed heuristic ranks the workers by the minimum level of incentives they need to receive in order to start working.

In our context, drivers always receive a guaranteed base pay when they work and sometimes they receive promotions on top of the base rate. We assume that the budget for promotions is separate from the budget for base rates. As not every driver who receives a promotion would choose to work, we compute two types of budgets for promotions. First, we compute the total promotions offered to all drivers for every shift on every day in the data as the projected budget. This is the total cost related to promotions incurred by the platform if all drivers chose to work. Second, we compute the actual cost based on the realized number of drivers who showed up to work at any given time as the realized budget. We can then compare the service capacity and cost of our heuristic relative to the actual allocation. As our data spans one year from October 2016 to September 2017, we choose the last nine months (January 1–September 30, 2017) as our test set. For each shift on each day in the test set, we train our model using all observations from the same shift and day of the week prior to the focal shift. Across 1,012 day-shifts, we observe that 94.59% of drivers were offered a promotion but only 18.4% of them activated the offer and chose to work. Moreover, 94% of the drivers who worked did not receive any promotion. These observations suggest that there is an opportunity to improve the current allocation of financial incentives.

To determine drivers' baseline probability to work, we first compute the average fraction of drivers who worked during a given shift on a given weekday using all past data, denoted by $\overline{D}$. We then compute the inverse c.d.f. evaluated at $\overline{D}$: $\tilde{D} = \Phi^{-1}(\overline{D})$, that is, $\tilde{D}$ represents the argument of $\Phi(\cdot)$ in the right hand-side of Equation (6). In other words, $\tilde{D}$ corresponds to the combination of drivers' attributes that will induce a probability of working equal to $\overline{D}$. For each driver i, we use all the covariates' values with the base pay (e.g., excluding promotions) in our fitted model. This will predict the probability of working when offered only the base rate, $\hat{p}_i^{base}$. If $\hat{p}_i^{base} \geq \tilde{D}$, we label the driver as "driving without promotion." For other drivers, we compute the difference, $\Delta_i = \tilde{D} - \hat{p}_i^{base} > 0$, to determine the level of additional incentive needed for them to work.

6.1.1. Improving Service Capacity While Keeping the Same Budget. Assuming that the platform has a fixed budget for promotions, we consider a strategy to recruit more workers under the same budget. We first determine the number of drivers who would work regardless of promotions (i.e., their base rates are appealing enough to motivate them to work), and then rank the remaining drivers by increasing values of Δ_i. We compute the minimum work-inducing promotion level by dividing Δ_i by the estimated coefficient $\hat{\beta}_{offer}$. We call this value $\tilde{\Delta}_i$. Then, a desired strategy is to allocate the promotion budget first to drivers with the smallest $\tilde{\Delta}_i$ until we exhaust the budget or we can no longer recruit additional drivers. On average, our proposed procedure sends promotions to 6.27% of all available drivers. The 95% interval for the fraction of drivers who should receive a promotion is [0.44%, 19.92%]; these fractions are substantially lower than the current practice of the company. As a result, a much smaller number of drivers would be targeted but each targeted driver would receive a much more attractive promotion.[4] Under the allocation observed in the data, drivers were offered an average promotion of 0.58× relative to their base rate. Under our proposed heuristic, however, targeted drivers receive an average promotion of 2.09×. Ultimately, using the same budget for promotions, our approach can staff 22.1% additional drivers on average with a 95% interval of [2.46%, 50.50%]. Figure 5 reports the percentage increase in the number of drivers for each shift and weekday.

6.1.2. Maintaining Service Capacity at a Lower Cost. Companies may have a target level of capacity they hope to meet for several reasons, such as to satisfy a high forecast demand or maintain low and reliable wait times. Similar to the previous case, we rank all drivers by increasing values of the minimum work-inducing promotion level (i.e., $\tilde{\Delta}_i$). We subtract the number of drivers who are predicted to work without any promotion from the desired service capacity. Instead of having a budget constraint, we now allocate promotions to drivers who require the smallest incentive $\tilde{\Delta}_i$ until we reach the desired service capacity. On average, the allocation under our heuristic costs 30.10% less relative to current practice with a 95% interval of [0.75%, 63.54%]. Figure 6 shows the percentage of cost savings for each shift and weekday.

6.2. Impact of Behavioral Explanations of Labor Decisions

In this section, we quantify the impact of capturing the main behavioral factors obtained in our estimation results. To this end, we investigate how many workers the platform would fail to attract if it did not incorporate income targeting and inertia into incentive design. We compare the following three scenarios to our model:

a. *ISF only*: The firm assumes that work decisions are influenced by *ISF* but not *HSF*.

b. *HSF only*: The firm assumes that work decisions are influenced by *HSF* but not *ISF*.

c. *Base*: The firm ignores both income-targeting and inertia behaviors.

Our analysis is at the day-shift level and reports out-of-sample predictions. The test set consists of each day-shift between January 1, 2017, and September 30, 2017. For each day-shift in the test set, we train four separate choice equations—one for each model (a)–(c) and one for our model—using all historical observations of the same day-shift from October 2016 to the week prior to the focal date. Each of the four choice equations represents the predicted outcome depending on the platform's assumption on workers' behavior. We first compute the fraction of drivers' work decisions that each model predicts correctly out-of-sample relative to the actual realization in the data. On average, our model outperforms the other three models in prediction accuracy both at the shift and day levels. Specifically, when the company ignores behavioral drivers of labor decisions, it loses 8.6% in prediction accuracy on average. Following the same procedure as in Section 6.1, we compute the incentive allocation under each model. More precisely, we first assume that each model is the true state of the world and solve for the optimal incentive allocation given the promotion budget observed in the data. Once the allocation is completed, we estimate the expected number of drivers who

Figure 5. (Color online) Number of Additional Drivers Using Our Allocation Strategy Given a Fixed Budget

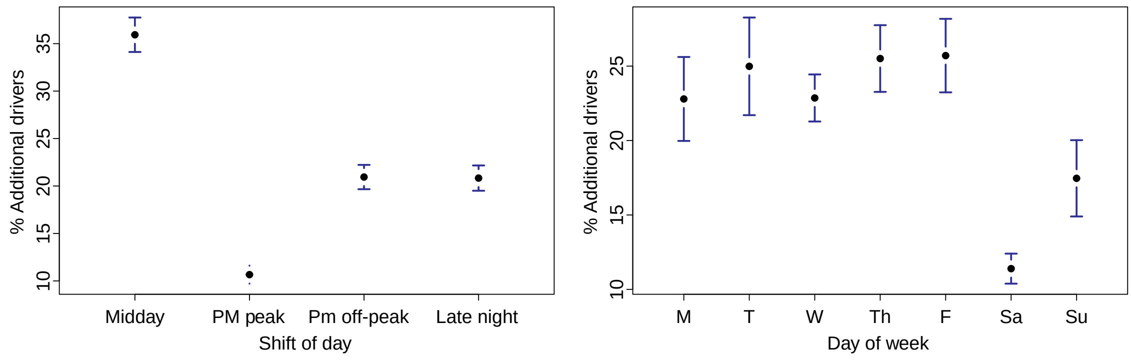
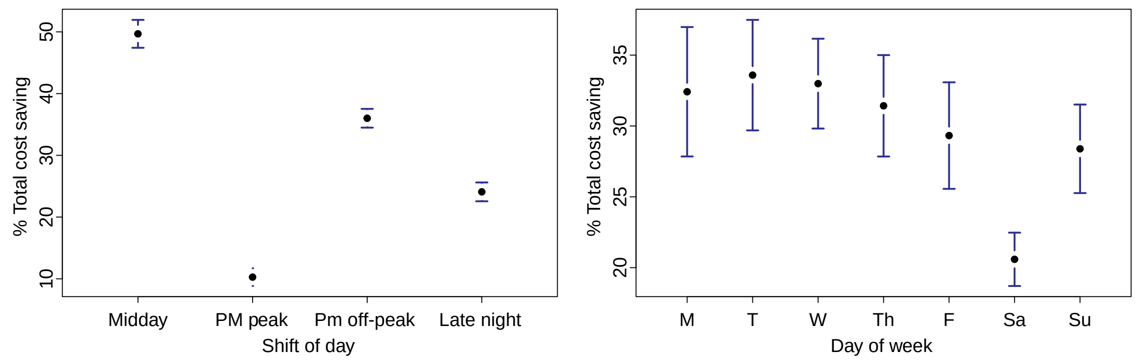


Figure 6. (Color online) Simulated Cost Savings While Maintaining the Same Service Capacity



would be working, assuming that the true state of the world is actually governed by our model. Note that by construction, our model will always outperform the other models in terms of expected capacity. Our main goal here is to quantify the magnitude of capacity loss when the company make different assumptions about workers' behavior.

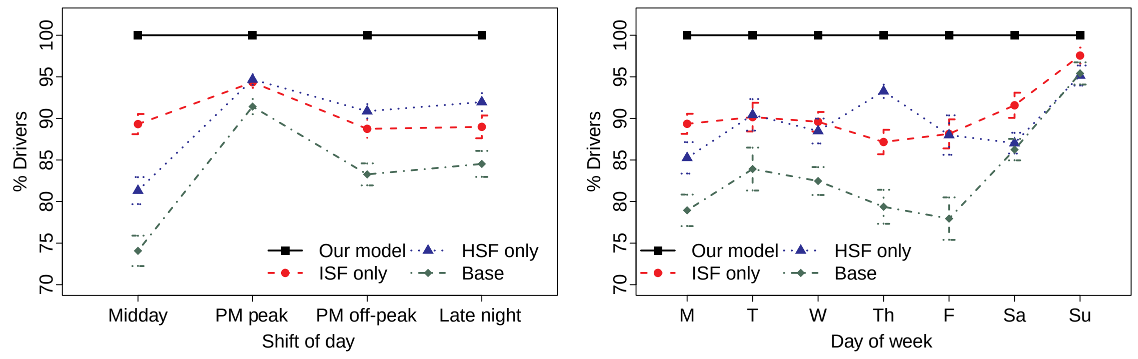
Figure 7 shows that ignoring behavioral factors can lead to a significant loss in the number of active drivers. Specifically, the base model leads to an average loss of 16.70% in the expected number of active drivers relative to our model, with a standard deviation of 13.06%. The *ISF* only (*HSF* only) model leads to an average reduction of 9.63% (10.32%) in the expected number of active drivers with a standard deviation of 9.10% (10.20%).

In summary, these results suggest that it is important for gig platforms to account for income targeting and inertia. Ignoring these behavioral motives can decrease prediction accuracy, and, more importantly, induce misleading incentive decisions that may result in suboptimal capacity levels.

7. Concluding Remarks

The recent rise of the gig economy has changed the way people think about employment. Unlike traditional employees who work under a fixed schedule, gig economy workers are free to choose their own schedule and platform to provide service. Such flexibility poses a great challenge to gig platforms in terms of planning and committing to a service capacity. It also poses a challenge to policymakers who are concerned about protecting workers. In this paper, we propose a framework to investigate how gig economy workers make labor decisions. Using data from a ride-hailing platform, we develop an econometric model that accounts for sample selection and endogeneity and controls for the competition within the ride-hailing industry. We find that financial incentives have a positive effect on the decision to work and on the work duration, confirming the positive income elasticity from the standard income effect. We also observe the influence of behavioral factors through the accumulated earnings and number of hours previously worked. The dominating effect, inertia, suggests that the longer workers have been working so far, the more likely they will continue working and the longer duration they will work for. Our results also reflect a unique feature of gig work. Whereas workers decide whether to work on both shift and day levels, they decide on work duration on a shift basis. Finally, our numerical experiments demonstrate that gig platforms can benefit from incorporating our insights into their incentive optimization.

Figure 7. (Color online) Impact of Ignoring Behavioral Factors on the Expected Number of Active Drivers



One of the important phenomena that emerge from this paper is the existence of inertia among drivers. Although we cannot conclude that all gig economy workers exhibit such a behavior, we believe that it has important implications that go beyond this study. Indeed, we believe that our findings are generalizable to other flexible workforces. Drivers in our data are not exclusive to the focal platform and are often working for other gig companies. Policies used by the focal platform are also quite common in the industry, from convenience delivery (e.g., GoPuff, Instacart) to local services (e.g., Rover, Handy) to tutoring services (e.g., Chegg Tutors, Magic Ears). Therefore, there is a lesson to be learned about the fundamental impact of such policies. Amid intensifying competition among providers of similar on-demand services, companies are making every effort to win over a mutual pool of workers. This paper empirically identifies several key behavioral factors that affect gig economy workers' decisions. These findings can be used to sharpen platforms' understanding on how gig economy workers make labor decisions and, ultimately, improve platforms' operational decisions (e.g., sending the right offer to the right worker at the right time).

This paper opens several avenues for future research. It could be interesting to validate our findings by running a controlled field experiment. Given that online platforms routinely run experiments to validate insights, testing the income targeting and inertia effects could be of interest. As our industry partner compensates workers using a dynamic hourly wage, future work could also compare the impact of behavioral and economic drivers under different compensation schemes. A second direction is to further investigate how workers construct their reference points or targets in both financial and time dimensions, and how these targets are updated over time. This will allow companies to gain insights about the (dis)utility of working as well as understanding how workers switch between service providers. With increasing data availability both within and across platforms, future researchers can address one of the limitations of this paper, namely the lack of granular data on each worker's decision, and extend our framework to deepen the analysis of gig economy workers' behaviors. Finally, our incentive allocation is based on simple ranking arguments. Developing a more comprehensive optimization framework to optimize incentives for each driver in each shift under further operational constraints and personalization is also an interesting extension. For example, it may be illegal or perceived as unfair to offer personalized wages at the worker level in certain regions so that the platform might prefer to stick to a limited menu of possible wage rates for their workers. As we have demonstrated the importance of workers' income targeting and inertia, researchers and practitioners could establish new policies based on these insights to improve the platform's operational outcome and ensure a high welfare for all parties. The main goals of this research stream would be to refine our understanding of gig economy workers and develop data-driven methods that can be used by gig economy platforms to efficiently motivate and strengthen their relationships with their flexible workforce.

Endnotes

[1] We cannot reveal the exact number of drivers and the size of our data set due to confidentiality. However, these exact numbers do not affect any of our results or findings.

[2] To ensure that drivers are not working for other platforms at the same time, the app will redirect idle drivers to a new waiting location every few minutes. Drivers have to confirm they reach the location via GPS.

[3] See https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page.

[4] One potential concern is fairness. Future research can include additional constraints such as the minimum fraction of drivers receiving a promotion and the maximum number of different promotion levels.

References

Agarwal S, Diao M, Pan J, Sing TF (2015) Are Singaporean cabdrivers target earners? Preprint, submitted October 11, 2013; revised March 27, 2015, https://dx.doi.org/10.2139/ssrn.2338476.

Beckers DG, van der Linden D, Smulders PG, Kompier MA, Taris TW, Geurts SA (2008) Voluntary or involuntary? Control over overtime and rewards for overtime in relation to fatigue and work satisfaction. *Work Stress* 22(1):33–50.

Benjaafar S, Ding JY, Kong G, Taylor T (2022) Labor welfare in on-demand service platforms. *Manufacturing Service Oper. Management* 24(1):110–124.

Benjaafar S, Hu M (2020) Operations management in the age of the sharing economy: What is old and what is new? *Manufacturing Service Oper. Management* 22(1):93–101.

Bourguignon F, Fournier M, Gurgand M (2007) Selection bias corrections based on the multinomial logit model: Monte Carlo comparisons. *J. Econom. Surveys* 21(1):174–205.

Brachet T, David G, Drechsler AM (2012) The effect of shift structure on performance. *Amer. Econom. J. Appl. Econom.* 4(2):219–246.

Bray RL, Serpa JC, Colak A (2019) Supply chain proximity and product quality. *Management Sci.* 65(9):4079–4099.

Cachon GP, Daniels KM, Lobel R (2017) The role of surge pricing on a service platform with self-scheduling capacity. *Manufacturing Service Oper. Management* 19(3):368–384.

Camerer C, Babcock L, Loewenstein G, Thaler R (1997) Labor supply of New York City cabdrivers: One day at a time. *Quart. J. Econom.* 112(2):407–441.

Carrington WJ (1996) The Alaskan labor market during the pipeline era. *J. Political Econom.* 104(1):186–218.

Chen MK (2016) Dynamic pricing in a labor market: Surge pricing and flexible work on the Uber platform. *Proc. 2016 ACM Conf. Econom. Comput.* (ACM, New York), 455.

Chen MK, Rossi PE, Chevalier JA, Oehlsen E (2019) The value of flexible work: Evidence from Uber drivers. *J. Political Econom.* 127(6):2735–2794.

Chen YJ, Dai T, Korpeoglu CG, Körpeoğlu E, Sahin O, Tang CS, Xiao S (2020) OM Forum–Innovative online platforms: Research opportunities. *Manufacturing Service Oper. Management* 22(3):430–445.

Collewet M, Sauermann J (2017) Working hours and productivity. *Labour Econom.* 47:96–106.

Corgnet B, Gaechter S, Hernán-González R (2020) Working too much for too little: Stochastic rewards cause work addiction. Preprint, submitted March 16, https://dx.doi.org/10.2139/ssrn.3540225.

Cragg JG (1971) Some statistical models for limited dependent variables with application to the demand for durable goods. Econometrica 39(5):829–844.

Crawford VP, Meng J (2011) New York City cab drivers' labor supply revisited: Reference-dependent preferences with rational-expectations targets for hours and income. *Amer. Econom. Rev.* 101(5):1912–1932.

Dahl GB (2002) Mobility and the return to education: Testing a Roy model with multiple markets. *Econometrica* 70(6): 2367–2420.

DeVoe SE, Lee BY, Pfeffer J (2010) Hourly vs. salaried payment and decisions about trading time and money over time. *Indust. Labor Relations Rev.* 63(4):627–640.

Dhaene G, Jochmans K (2015) Split-panel jackknife estimation of fixed-effect models. *Rev. Econom. Stud.* 82(3):991–1030.

Dong J, Ibrahim R (2020) Managing supply in the on-demand economy: Flexible workers, full-time employees, or both? *Oper. Res.* 68(4):1238–1264.

Donohue K, Özer Ö, Zheng Y (2020) Behavioral operations: Past, present, and future. *Manufacturing Service Oper. Management* 22(1):191–202.

Farber HS (2005) Is tomorrow another day? The labor supply of New York City cabdrivers. *J. Political Econom.* 113(1):46–82.

Farber HS (2008) Reference-dependent preferences and labor supply: The case of New York City taxi drivers. *Amer. Econom. Rev.* 98(3):1069–1082.

Farber HS (2015) Why you can't find a taxi in the rain and other labor supply lessons from cab drivers. *Quart. J. Econom.* 130(4): 1975–2026.

Farewell V, Long D, Tom B, Yiu S, Su L (2017) Two-part and related regression models for longitudinal data. *Annual Rev. Statist. Appl.* 4:283–315.

Greene W (2004) Fixed effects and bias due to the incidental parameters problem in the tobit model. *Econometric Rev.* 23(2):125–147.

Hahn J, Newey W (2004) Jackknife and analytical bias reduction for nonlinear panel models. *Econometrica* 72(4):1295–1319.

Heckman JJ (1979) Sample selection bias as a specification error. *Econometrica* 47:153–161.

Hu M, Zhou Y (2020) Price, wage, and fixed commission in on-demand matching. Preprint, submitted April 10, 2017; last revised September 1, 2020, https://dx.doi.org/10.2139/ssrn.2949513.

Ibrahim R (2018) Managing queueing systems where capacity is random and customers are impatient. *Production Oper. Management* 27(2):234–250.

Imbens G, Wooldridge J (2007) Control function and related methods, what's new in econometrics (National Bureau of Economic Research, Cambridge, MA). https://users.nber.org/~confer/2007/si2007/WNE/lect_6_controlfuncs.pdf.

Karacaoglu N, Moreno A, Ozkan C (2018) Strategically giving service: The effect of real-time information on service efficiency. Preprint submitted October 27, https://dx.doi.org/10.2139/ssrn.3260035.

Leclerc F, Schmitt BH, Dube L (1995) Waiting time and decision making: Is time like money? *J. Consumer Res.* 22(1):110–119.

Levinsohn J, Petrin A (2003) Estimating production functions using inputs to control for unobservables. *Rev. Econom. Stud.* 70(2): 317–341.

Madden D (2008) Sample selection vs. two-part models revisited: The case of female smoking and drinking. *J. Health Econom.* 27(2): 300–307.

Manyika J, Lund S, Robinson K, Valentino J, Dobbs R (2015) A labor market that works: Connecting talent with opportunity in the digital age. Report, McKinsey Global Institute, Washington, DC. https://www.mckinsey.com/featured-insights/employment-and-growth/connecting-talent-with-opportunity-in-the-digital-age.

Mason CH, Perreault WD Jr (1991) Collinearity, power, and interpretation of multiple regression analysis. *J. Marketing Res.* 28(3):268–280.

Newey WK (1987) Efficient estimation of limited dependent variable models with endogenous explanatory variables. *J. Econometrics* 36(3):231–250.

Oettinger GS (1999) An empirical analysis of the daily labor supply of stadium vendors. *J. Political Econom.* 107(2):360–392.

Ozimek A (2021) Freelance forward economist report. Report, Economic Innovation Group. https://www.upwork.com/research/freelance-forward-2021.

Puhani P (2000) The Heckman correction for sample selection and its critique. *J. Econom. Surveys* 14(1):53–68.

Sheldon M (2016) Income targeting and the ridesharing market. Working paper, http://www.michaelsheldon.org/working-papers.

Soman D (2001) The mental accounting of sunk time costs: Why time is not like money. *J. Behav. Decision Making* 14(3):169–185.

Stafford TM (2015) What do fishermen tell us that taxi drivers do not? An empirical investigation of labor supply. *J. Labor Econom.* 33(3):683–710.

Taylor TA (2018) On-demand service platforms. *Manufacturing Service Oper. Management* 20(4):704–720.

Thakral N, Tô LT (2021) Daily labor supply and adaptive reference points. *Amer. Econom. Rev.* 111(8):2417–2443.

Wooldridge JM (2015) Control function methods in applied econometrics. *J. Human Resources* 50(2):420–445.