2021

Managing The Gig Economy Via Behavioral And Operational Lenses

Wichinpong Park Sinchaisri
University of Pennsylvania

Managing The Gig Economy Via Behavioral And Operational Lenses

Abstract

This dissertation combines tools from operations management, econometrics, machine learning, and behavioral sciences to (i) study how on-demand workers learn and make decisions in complex environments, (ii) develop tools to help improve their decision-making, and (iii) inform the design of better policies to manage human-centered operations. Recent technologies create and accelerate new work arrangements that provide workers with flexibility in their schedule and choice of service. At the same time, the decisions a worker faces have become more complex. Platforms offer competing dynamic incentives, and the independent nature of gig work means that workers do not experience the benefits of learning from colleagues. The following three chapters investigate how behavioral operations management can be utilized to better manage the gig economy.

(i) Behavioral and Economic Drivers of Decisions.We empirically investigate how on-demand workers decide on when to work and \emph{for how long} depending on varying financial incentives and personal goals. Using the comprehensive data from a ride-hailing industry partner, we develop an econometric framework that addresses empirical challenges such as sample selection bias and endogeneity. Our results demonstrate that, while workers exhibit positive income elasticity as predicted by standard income theory, their decisions are significantly influenced by their cumulative earnings (more likely to stop working when reaching their income goal) and recent work duration (tend to stay working after long hours of work or exhibit "inertia"), more akin to the behavioral theory of labor supply. Inertia captures both the formation of work habits and the tendency to stay with the focal platform, suggesting that, amidst intensifying competition among platforms, platform loyalty could be induced through optimal incentive design. Thus, we propose a heuristic to optimize incentive allocation and demonstrate through counterfactual simulations the monetary and capacity benefits of accounting for our behavioral insights.

(ii) Dynamic Decisions and Multihoming Behavior. We leverage proprietary data from our ride-hailing industry partner and the publicly available trip record data to develop and estimate a structural behavioral model of gig workers' sequential dynamic decisions of when and where to work in the presence of alternative work opportunities. Our major contributions are in the modeling and estimation of dynamic decisions with temporal and spatial components and dynamic outside options, and the development of an efficient simulation-assisted machine learning-based estimation framework. Our results characterize gig workers' forward-looking behavior and heterogeneous cost of working. We find that workers are strategic in their choice of initial service location to ensure high utilization and are prone to multihoming behavior when facing longer idle times. Then, we study how the firm can influence multihoming behavior among workers. Our counterfactual analyses demonstrate the effectiveness of strategies commonly used in practice and offer insights that can help retain workers during high demand or nudge them to quit during low demand.

(iii) Improving Human Decision-Making with Machine Learning. We propose a novel machine-learning algorithm to automatically extract best practices from the trace data and infer simple tips that can help workers learn to make better decisions. We use an approach based on imitation learning and interpretable reinforcement learning and consider simple if-then-else rules that modify workers' strategy in a way that most improve their performance, capture useful insights that are challenging for workers to learn by themselves, and are simple enough for workers to understand. To validate our approach and test the performance of our algorithm, we design a virtual kitchen-management game and conduct large-scale pre-registered behavioral studies on Amazon Mechanical Turk. Our experiments show that rules inferred from our algorithm are effective and significantly outperform rules from other sources at improving performance and speeding up learning among workers. In particular, we help workers identify optimal early actions that help them improve in the long term and discover additional optimal strategies beyond

what is stated by our algorithm.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Operations & Information Management

First Advisor
Gad Allon

Keywords
Behavioral Operations Management, Empirical Operations Management, Future of Work, Gig Economy, Human-AI Interface, Incentives

Subject Categories
Business Administration, Management, and Operations | Economics | Management Sciences and Quantitative Methods | Operational Research

MANAGING THE GIG ECONOMY
VIA BEHAVIORAL AND OPERATIONAL LENSES

Wichinpong Park Sinchaisri

A DISSERTATION

in

Operations, Information, and Decisions Department

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2021

Supervisor of Dissertation

Gad Allon, Jeffrey A. Keswin Professor of Operations, Information, and Decisions

Graduate Group Chairperson

Nancy Zhang, Ge Li and Ning Zhao Professor, Professor of Statistics

Dissertation Committee

Hamsa Bastani, Assistant Professor of Operations, Information, and Decisions

Ken Moon, Assistant Professor of Operations, Information, and Decisions

Serguei Netessine, Dhirubhai Ambani Professor of Innovation and Entrepreneurship, Professor of Operations, Information, and Decisions, Vice Dean for Global Initiatives

MANAGING THE GIG ECONOMY

VIA BEHAVIORAL AND OPERATIONAL LENSES

Dedicated to my parents who have long instilled the passion for OM in me.

ACKNOWLEDGEMENT

I would like to express my deepest appreciation to my dissertation committee. First, I am extremely grateful to my PhD advisor, Gad Allon, for his guidance, support, and wisdom over the past five years. He has taught me tremendously how to be a great thinker, researcher, teacher, communicator, and an all-around better person. Thank you so much, Gad, for believing in me, encouraging and pushing me to succeed, and giving me opportunities to learn and grow. I am also deeply indebted to Serguei Netessine, Ken Moon, and Hamsa Bastani. Serguei, thank you for your great kindness and support, from a research collaboration to the job market. Ken, thank you for always inspiring me with cool research ideas and methods and for helping me navigate my research journey since the first semester of my PhD. Hamsa, thank you for making research really fun and for inspiring me to combine my behavioral OM passion with machine learning. Working with you and Osbert has helped shape and expand my vision for future research.

I would like to extend my sincere thanks to other faculty members in the Operations, Information, and Decisions Department who have inspired, supported, or taught me in various occasions: Noah Gans, Etan Green, Cade Massey, Katherine Milkman, Alice Moon, Maria Rieders, Sergei Savin, Rebecca Schaumberg, Uri Simonsohn, Hummy Song, Xuanming Su, Christian Terwiesch, and Senthil Veeraraghavan. I also wish to thank other Penn faculty members: Osbert Bastani, Lindsey Cameron, Ulrich Doraszelski, Stephanie Doupnik, Martha Farah, Gideon Nave, Ingrid Nembhard, Dylan Small, and Christoph van De Bulte. I would like to thank Shane Jensen for advising my dual master's thesis and introducing me to urban analytics research. I would like to thank Maxime Cohen for proving advice and support as well as connecting us with our industry partner. I must also thank my Master's co-advisors, Georgia Perakis and Yanchong Karen Zheng, for inspiring me to pursue a PhD in OM and for their continued support and advice. Finally, special thanks to the teams at Wharton OIDD, Wharton Doctoral Programs, Wharton Behavioral Lab, Wharton Computing, and Wharton Social Impact Initiative.

I have been able to experience the fun side of PhD thanks to the amazing friends I have made throughout the past five years. I am grateful to the seniors who have helped and inspired me with their wholesomeness: Edward Chang, Youran Fu, Rowena Gan, Fei Gao, Chloe Kim Glaeser, Joshua Lewis, Dawson Kaaua, Lesley Meng, Alex Miller, Vanitha Virudachalam, Ozge Yapar, and Jiding Zhang. Many thanks to my fellow OIDD-ers: Arielle Anderer, Daniel Chen, Christian Kaps, Erika Kirgios, Tan Lekwijit, Katie Mehr, Aneesh Rai, Pia Ramchandani, Amandeep Singh, Sam Skowronek, Canberk Ucel, and Sirui Wang, and my Penn/Wharton friends: Allison Briggs, Junhui Cai, Jaeho Choi, Andrea Contigiani, Molly Frean, Ellen Fu, Nathan Guo, Lauren Kaufmann, Joowon Klusowski, Jong Myeong Lim, Yi Liu, Joseph Martinez, and Stella Park. Warm thanks to the Penn Thai community, including Pim Hannanta-anan, Tee Ramaithitima, Paclink Thaweerattanasinp, Fair Pisuttisarun, Toffy Charupatanapongse, Nat Puapattanakajorn, and Dante Lokitiyakul. I would also like to thank Philippe Blaettchen and Nazlı Sömnez, my awesome job market buddies, as well as Kashish Arora and MoonSoo Choi for making going through the job market during the COVID-19 pandemic memorable.

Another major part of my experience at Penn is being part of Stouffer College House. For four years, I have received so much love and support from and made lifelong friendships and memories with the Stouffer community. I am extremely grateful to Nadir Sharif and Phil Nichols for this precious opportunity for me to grow as a leader and foster a wonderful community. I would also like to thank my colleagues and residents, including Lin Yuhan, Yun Cha, Chenelle Boatswain, Yoonhee Ha, Xiong Her, Paola Pérez Moreno, Tahira Islam, and my B Section/Mayer 2 folks. Nothing *really* comes closer to home.

Finally, I would like to thank and dedicate this dissertation to my parents. They have been my biggest champion and their constant encouragement always got me through the good and bad times. It was them who first taught me about decision-making and fueled my curiosity about optimization. From the kid who loved observing queues at transit stations and theme parks, I could not have come this far without their love, support, and inspiration.

ABSTRACT

MANAGING THE GIG ECONOMY
VIA BEHAVIORAL AND OPERATIONAL LENSES

Wichinpong Park Sinchaisri

Gad Allon

This dissertation combines tools from operations management, econometrics, machine learning, and behavioral sciences to (i) study how gig economy workers learn and make decisions in complex environments, (ii) develop tools to help improve their decision-making, and (iii) inform the design of better policies to manage human-centered operations. Recent technologies create and accelerate new work arrangements that provide workers with flexibility in their schedule and choice of service. At the same time, the decisions a worker faces have become more complex. Platforms offer competing dynamic incentives, and the independent nature of gig work means that workers do not experience the benefits of learning from colleagues. The following three chapters investigate how behavioral operations management can offer insights into how to better manage the flexible workforce.

(i) Behavioral and Economic Drivers of Decisions. We empirically investigate how on-demand workers decide on *when to work* and *for how long* depending on varying financial incentives and personal goals. Using the comprehensive data from a ride-hailing industry partner, we develop an econometric framework that addresses empirical challenges such as sample selection bias and endogeneity. Our results demonstrate that, while workers exhibit positive income elasticity as predicted by standard income theory, their decisions are significantly influenced by their cumulative earnings (more likely to stop working when reaching their income goal) and recent work duration (tend to stay working after long hours of work or exhibit *"inertia"*), more akin to the behavioral theory of labor supply. Inertia captures both the formation of work habits and the tendency to stay with the focal platform, suggesting that, amidst intensifying competition among platforms, platform loyalty could

be induced through optimal incentive design. Thus, we propose a heuristic to optimize incentive allocation and demonstrate through counterfactual simulations the monetary and capacity benefits of accounting for our behavioral insights.

(ii) Dynamic Decisions and Multihoming Behavior. We leverage proprietary data from our ride-hailing industry partner and the publicly available trip record data to develop and estimate a structural model of gig workers' sequential dynamic decisions of *when* and *where* to work. Our major contributions are in the modeling and estimation of dynamic decisions with temporal and spatial components and dynamic outside options, and the development of an efficient simulation-assisted machine learning-based estimation framework. Our results characterize workers' forward-looking behavior and the distribution of operating cost. We find that workers are strategic in their choice of initial service location to ensure high utilization and are prone to multihoming when facing longer idle times. Then, we study how the firm can control workers' multihoming behavior. Our counterfactual analyses demonstrate the effectiveness of strategies commonly used in practice and offer insights that can help retain workers during high demand or nudge them to quit during low demand.

(iii) Improving Human Decision-Making with Machine Learning. We propose a novel machine-learning algorithm to automatically extract best practices from the trace data and infer simple tips that can help workers learn to make better decisions. We use an approach based on imitation learning and interpretable reinforcement learning and consider simple if-then-else rules that modify workers' strategy in a way that most improve their performance, capture useful insights that are challenging for workers to learn by themselves, and are simple enough for workers to understand. To validate our approach and test the performance of our algorithm, we design a virtual kitchen-management game and conduct large-scale pre-registered behavioral studies on Amazon Mechanical Turk. Our experiments show that rules inferred from our algorithm are effective and significantly outperform rules from other sources at improving performance and speeding up learning among workers. In particular, we help workers identify optimal early actions that help them improve in the long term and discover additional optimal strategies beyond what is stated by our algorithm.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1 : The Impact of Behavioral and Economic Drivers on Gig Economy Workers

1.1. 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 2019, 57 million Americans or 35% of the U.S. workforce engaged in gig work (Intelligence 2019), providing a wide range of services, from ride-hailing (e.g., Uber, Lyft) to food delivery (e.g., DoorDash, Caviar) 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 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. 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 2019). 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., Benjaafar et al. 2019, Dong and Ibrahim 2020), 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 (Donohue et al. 2020, Chen et al. 2020).

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 are the factors that 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 dataset 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 incentives and conduct counterfactual simulations to examine its performance and quantify potential losses if the company ignores workers' behavior when designing incentives.

Contributions.

Our paper contributes to the economics and operations literatures in four ways. First, we offer a potential way to reconcile the two competing 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 on 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. Second, we unravel a new behavioral driver of labor decisions, *inertia*. 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 or inertia. Furthermore, it can potentially hint at workers' loyalty to the focal platform. 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 their incentive policy accounting for workers' behavior, it can increase the capacity by 22% without incurring additional cost or maintain the same service level at a 30% lower cost.

1.2. 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.

1.2.1. 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.

1.2.2. 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) and Farber (2008) find 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 conclude that the findings are not robust. Crawford and Meng (2011) implements 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 (2019) estimates 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.

1.2.3. 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 non-monetary 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. While 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 and Sheldon 2016) and benefit drivers via better utilization (Cachon et al. 2017). Hu and Zhou (2019) studies 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, Benjaafar et al. 2019, Dong and Ibrahim 2020).

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) studies e-hailing taxi drivers in South America and finds that drivers' response to real-time information about other drivers ' locations could explain different utilization they can achieve. Chen et al. (2019) documents how Uber drivers value real-time flexibility and estimates 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. While 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 datasets record only the trips that happened. In our dataset, 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.

1.2.4. 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. Since individuals' targets cannot be observed, we use workers' accumulated earnings since the beginning of their work day 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.

H1: A higher wage increases the probability of working and the work duration.

Following the standard income effect (see §1.2.1), 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 H2 and H3 below). 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.

H2: Higher accumulated earnings decreases the probability of working and the work duration.

Studies of taxi drivers including Camerer et al. (1997), Farber (2008), and Thakral and Tô (2019) provide support for an income-targeting behavior; that is, the probability to stop working increases once the income target is reached. Thakral and Tô (2019) further demonstrates 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 towards 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 on the work duration.

H3: 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) develops 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 H2, 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.

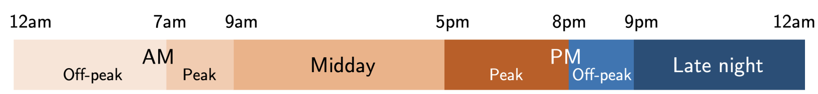
1.3. Data: Ride-hailing Platform in New York City

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 dataset 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.

1.3.1. 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. Then, the platform will match 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.

Figure 1: Breakdown of shifts for each operating day



1.3.2. Shifts and Work Schedule

Each operating day is divided into six shifts specified by the company (see an illustration in Figure 1): morning non-rush hours from midnight to 7am (*AM Off-peak*), morning rush hours from 7 to 9am (*AM Peak*), midday from 9am to 5pm (*Midday*), afternoon rush hours from 5 to 8pm (*PM Peak*), evening non-rush hours from 8 to 9pm (*PM Off-peak*), and late night from 9pm to midnight (*Late night*). The largest volume of activities happen during PM Off-peak, followed by PM Peak, and Midday, while 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

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

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, while weekend trips account for 19.62%. While 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. 91.07% of drivers' working days did not overlap with midnight (e.g., they did not work overnight).

1.3.3. 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 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×.

[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.

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 pro-rated 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 eranings to their bank account.

1.3.4. 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 dataset, we record both their first day joining the platform and their first work day 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. These allow 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 3-passenger sedan, a small 3-passenger SUV, a medium 4-passenger SUV, a large 5-passenger SUV, a 5-passenger van, and a 6-passenger van. We exclude van drivers from our analysis as the majority of them lease their vehicle from the company rather than owning their vehicle or leasing 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, while 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.

1.3.5. 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 have 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 §1.4.1, we discuss the metrics that we construct to control for market conditions and competition intensity.

1.4. Empirical Approach

To test the hypotheses developed in §1.2, 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,

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

then provide details of our estimation method, and finally discuss the empirical challenges and our strategies to address them.

1.4.1. Empirical Model and Estimation Details

As discussed, our dataset provides a unique advantage as we observe the financial incentives offered to *every* driver for *every* shift as long as the driver 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 §1.4.2 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 OLS regression.

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 5 minutes. We also exclude public holidays from our analysis.

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 the driver starts 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 2019). *(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.

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.

$$Hour_{i,t} = \begin{cases} Hour^*_{i,t} & \text{if} \quad Drive_{i,t} = 1 \\ \text{unobserved} & \text{otherwise} \end{cases} \tag{1.1}$$

$$Drive_{i,t} = \begin{cases} 1 \quad \text{if} \quad Drive^*_{i,t} > 0 \\ 0 \quad \text{otherwise} \end{cases} \tag{1.2}$$

$$Drive^*_{i,t} = \alpha_{0,i} + \alpha_w w_{i,t} + \alpha_{ISF} ISF_{i,t} + \alpha_{HSF} HSF_{i,t} + \boldsymbol{\alpha X_{i,t}} + v_{i,t} \tag{1.3}$$

$$Hour^*_{i,t} = \beta_{0,i} + \beta_{\tilde{w}} \tilde{w}_{i,t} + \beta_{ISF} ISF_{i,t} + \beta_{HSF} HSF_{i,t} + \boldsymbol{\beta Z_{i,t}} + u_{i,t} \tag{1.4}$$

$$\begin{bmatrix} \sigma_v^2 \\ \sigma_u^2 \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho\sigma_u \\ \rho\sigma_u & \sigma_u^2 \end{bmatrix} \right). \tag{1.5}$$

The two stages that we estimate are given by:

$$P(Drive_{i,t} = 1 | \boldsymbol{X_{i,t}}) = \Phi(\alpha_{0,i} + \alpha_w w_{i,t} + \alpha_{ISF} ISF_{i,t} + \alpha_{HSF} HSF_{i,t} + \boldsymbol{\alpha X_{i,t}}), \tag{1.6}$$

$$f(Hour_{i,t}) = \beta_{0,i} + \beta_{\tilde{w}} \tilde{w}_{i,t} + \beta_{ISF} ISF_{i,t} + \beta_{HSF} HSF_{i,t} + \boldsymbol{\beta Z_{i,t}} + \theta\lambda_{i,t} + u_{i,t}, \tag{1.7}$$

where $\Phi(\cdot)$ is the normal c.d.f. and $\lambda_{i,t}$ is the inverse Mills ratio (IMR) calculated from the predicted probability in Equation (1.6) (*"Choice Equation"*). Thus, we essentially estimate a probit model for the work decision in Equation (1.6) 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 (1.7)), 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.

***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 §1.4.2). 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.

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.

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, 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, number of other ride-hailing trips during the same shift or day (in thousands).

1.4.2. Empirical Challenges and Strategies

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 which 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. Since we observe incentives for all drivers on every shift regardless of their work decisions, we can directly estimate the selection problem. As presented in §1.4.1, we employ a modified two-stage Heckman estimation method for our analysis.

While 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 §1.4.1) 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 §3.5. Finally, as a robustness check, we consider the 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) that provides Monte Carlo comparisons across different selection models and to Bray et al. (2019) that 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 Appendix B.1.

Endogeneity.

As discussed in §1.2.1, 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 dataset, 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).

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 *co-workers*. For each driver who is available to work at a particular time (i.e., has not terminated their partnership with the platform), we define their co-workers 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 co-workers for the focal period as an IV. This IV satisfies the *relevance condition:* since both the focal driver and their co-workers 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 co-workers 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 co-workers' 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 *co-skippers*. 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 Appendix B.2.

Multicollinearity.

A potential concern of including both HSF and ISF in the same specification is the multicollinearity issue. Correlations between ISF and HSF in our data range between 0.446

and 0.929, 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, while 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 §3.5) 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 Jr 1991). Our dataset 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 considering models with only ISF or HSF, performing localized regressions by controlling for drivers with similar ISF or HSF, and converting one of the two variables to be categorical. Our insights remain qualitatively consistent. Further details and discussion are deferred to Appendix B.3.

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 dataset, 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 reflects 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 to 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 Appendix C.

1.5. 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. The results for the Midday shift are discussed in detail and a summary of results for the remaining shifts is subsequently provided. 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 §1.2. Finally, we conduct several robustness tests that help validate our findings.

1.5.1. Within-Day Analysis

We examine drivers' labor decisions at the beginning of each of the company-specified shifts as introduced in §1.3.2. As 91% of drivers' working days observed in our data do not overlap with midnight and 73% of work day happened between 7am and midnight, we assume that the first possible shift of the day is AM Peak (starting at 7am) and the last possible 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 everyday after midnight.

For each shift, we first estimate the choice equation (Equation (1.6)) in which the outcome variable is a binary decision of whether to work for the focal shift. We then estimate the level equation (Equation (1.7)) that concerns the work duration for the shift, conditional on the decision to work. We compare three model specifications for the second stage: (i) baseline OLS, (ii) 2SLS without correction for sample selection bias ("two-part model"), and (iii) our main model which is a 2SLS with sample selection correction. Tables 1 and 2 display our estimates for the Midday shift of SUV and sedan drivers, respectively. The first column in each table reports the estimates from the control function probit of the choice equation. The second column reports the estimates from the baseline OLS for the level equation replicating the model implemented in the literature (Camerer et al. 1997, Sheldon 2016). We follow the model specification and IV strategy used in past work. Covariates include log hourly wage, temperature, rain indicator, day of week, and month dummies and we use the average of other drivers' hourly wages as an instrument. We then present the estimates from the level equation of the two-part model in the third column, and from the level equation of our main model in the fourth column.

Table 1: Estimates of two-stage selection models of SUV drivers' decisions during Midday shifts

	Choice Eq	*Level Eq Baseline*	*Level Eq Two-Part*	*Level Eq Main Model*
Incentives/targets				
Offer/Earnings	0.002*** (0.0006)	−0.083*** (0.019)	0.001 (0.001)	0.001 (0.001)
Income so far	−0.017*** (0.004)	-	−0.009*** (0.002)	−0.008*** (0.002)
Hours so far	2.904*** (0.163)	-	1.690*** (0.068)	1.826*** (0.070)
Hours last week				
Total	0.017*** (0.0003)	-	-	-
Same shift	-	-	0.056*** (0.002)	0.059*** (0.002)
New driver	0.590*** (0.060)	-	-	-
IMR	-	-	-	0.271*** (0.029))
Observations	124,769	45,330	45,329	45,329
R^2	-	0.378	0.552	0.552
Note:				*p<0.05; **p<0.01; ***p<0.001

SUV drivers.

For the choice equation, we find that hourly financial offer and cumulative work hours have a significantly positive impact on the decision to work, while cumulative earnings have a significantly negative impact. The first effect indicates that drivers respond positively to an increase in financial incentives as predicted by the standard income effect. The positive effect of HSF suggests that drivers who have worked for a longer period of time during the preceding shift (e.g., AM Peak), controlling for other covariates, are more likely to work for a new shift (e.g., Midday). We refer to this behavior as *inertia*, which we will discuss further as it becomes more prevalent across different analyses. In contrast, the negative effect of ISF reflects a potential income-targeting behavior, that is, drivers are less likely to work if they have earned more income or become closer to their (unobserved) income target. We also find that the number of hours each driver worked in the previous week has a significant positive impact on the decision to work. This could suggest that drivers tend to stick to their work patterns and hold relatively stable work schedules, as observed in Chen et al. (2019). In other words, past work decisions could play an important role in how drivers form and adjust their income and time targets. Lastly, we observe that newer drivers who recently joined the platform are significantly more likely to work.

We next consider the level equation of work duration. Interestingly, under the baseline model, we observe that SUV drivers exhibit a negative income elasticity, similar to full-time taxi drivers investigated in Camerer et al. (1997) and Thakral and Tô (2019), rather than a positive income elasticity observed for ride-hailing drivers (Sheldon 2016). For the other two models in which we incorporate proxies for income and time targets, the estimates for the level equation are relatively consistent regardless of sample selection correction. We observe a directional positive impact of hourly earnings on work duration, providing additional evidence that drivers exhibit positive income elasticity. The impact of ISF is significantly negative, suggesting that income-targeting behavior also negatively affects work duration. On the other hand, the impact of HSF or inertia behavior is significantly positive. We again observe that drivers might stick to their schedules as the work duration for the focal shift is positively affected by the work duration during the same shift in the previous week. In addition, the estimated coefficient of our sample selection correction variable (IMR) is statistically significant, confirming that selection into working is not random. Overall, we observe that the positive effects of hourly earnings and HSF dominate the negative impact of ISF on the work duration.

Sedan drivers.

We perform the same estimation and obtain similar results for sedan drivers: hourly offer or earnings rate and HSF have a positive impact on the decision to work and on the work

Table 2: Estimates of two-stage selection models of sedan drivers' decisions during Midday shifts

	Choice Eq	*Level Eq Baseline*	*Level Eq Two-Part*	*Level Eq Main Model*
Incentives/targets				
Offer/Earnings	0.007*** (0.0008)	0.080*** (0.028)	0.001 (0.001)	0.001 (0.001)
Income so far	−0.031*** (0.006)	-	−0.007*** (0.002)	−0.007*** (0.002)
Hours so far	3.243*** (0.192)	-	1.073*** (0.058)	1.058*** (0.061)
Hours last week				
Total	0.022*** (0.0004)	-	-	-
Same shift	-	-	0.079*** (0.003)	0.078*** (0.003)
New driver	0.660*** (0.042)	-	-	-
IMR	-	-	-	-0.029 (0.029)
Observations	113,444	20,307	20,297	20,297
R^2	-	0.389	0.580	0.580
Note:				*p<0.05; **p<0.01; ***p<0.001

duration. Under the baseline approach, we observe that, for sedan drivers, (log) hourly earnings rate positively affects the number of hours worked. The positive income elasticity is in line with findings from ride-hailing drivers in Sheldon (2016). This may suggest that SUV and sedan drivers are fundamentally different types of workers: SUV drivers' behaviors are similar to full-time professional taxi drivers, whereas sedan drivers' behaviors are similar to average drivers on ride-hailing platforms. While descriptive statistics suggest that SUV drivers tend to drive more often and for longer periods relative to sedan drivers, both types of drivers exhibit similar responses to hourly incentive, cumulative earnings, and work hours. Note that the estimated coefficient for IMR is not statistically significant (at $p = 0.05$) for this shift, suggesting that the evidence of selection of bias is weak. Nevertheless, our insights remain valid as the estimates are consistent regardless of sample selection correction. Furthermore, IMR estimates are statistically significant for all the other shifts (see Appendix A).

Estimates for other shifts.

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 contains the sign of the effect ($+$ or $-$) and its statistical significance

Figure 2: Signs and statistical significance for estimates of two-stage models of drivers' shift-level decisions

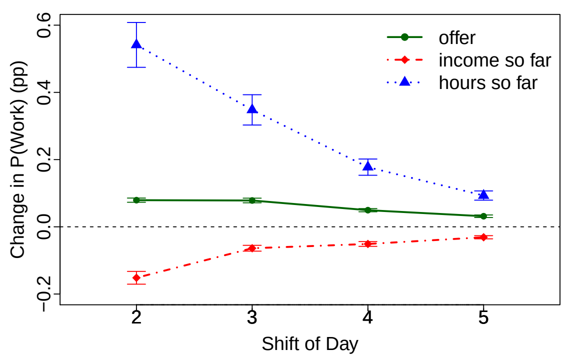
	Choice (Work or not)						Level (How long)						
SUV	Mean	IV-F	Offer	ISF	HSF	N	Mean	IV-F	Earn	ISF	HSF	R^2	N
Midday	0.343	372.9	**+**	**-**	**+**	124,769	4.987	180.2	*+*	**-**	**+**	0.552	45,329
PM-Peak	0.277	345.1	**+**	**-**	**+**	131,910	2.421	58.5	**+**	**-**	**+**	0.244	39,592
PM-OPeak	0.182	320.6	**+**	**-**	**+**	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-OPeak	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

Note: Solid background with bolded +: significantly positive, striped with bolded -: significantly negative, white with italicized sign: non-significant. All at $p = 0.05$.

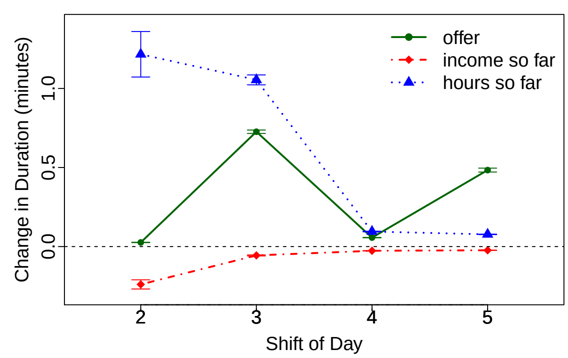
at $p = 0.05$ as follows: solid background with a bolded + indicates a significant positive estimate, striped background with a bolded − indicates a significant negative estimate, and white background with italicized sign corresponds to a non-significant directional effect. In addition, we provide the 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 observe that the estimates for drivers of both vehicle types are substantially 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, that is, financial incentives encourage the decision to work and boosts the work duration. However, we also observe an evidence of behavioral factors of labor supply with regards 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 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 fairly 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.

Figure 3: 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

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 one percent increase in each of the variables for an average driver. Figures 3a and 3b illustrate the change in probability of 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 on in the day. The detailed effect sizes for both SUV and sedan drivers are reported in Appendix A.

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.*

1.5.2. 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 progresses 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 §1.5.1, 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 vehicles types.

Figure 4: 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

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

At a day level, we draw considerably different conclusions from our shift-level analysis. While the positive impact of HSF on a decision to work remains consistent, the impact of hourly offer and ISF appear to vary across different 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 on 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, while 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.

1.5.3. 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 3 summarizes our hypotheses and results. We find that, as predicted by the standard income effect, drivers respond positively to financial incentives. While 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 have 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 cumulative income rather than the hourly wage. Thakral and Tô (2019) similarly demonstrates the existence of income targeting among taxi drivers and identifies the recently earned cumulative income as a key factor in the decision to quit.

In addition, we establish a new behavioral phenomenon. Workers who have previously

Table 3: Summary of hypotheses and results

		Shift-level		Day-level	
	Statement	SUV	Sedan	SUV	Sedan
H1a	Higher wage increases P(work)	✓	✓	✓→ ✗	✓→ ✗
H1b	Higher wage increases work hours	✓	✓	✗	✗
H2a	Higher income so far decreases P(work)	✓	✓	✗→ ✓	✗→ ✓
H2a	Higher income so far shortens work duration	✓	✓	✗	✗
H3a	Longer work hours so far decreases P(work)	✗	✗	✗	✗
H3b	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.

worked for a longer duration are more likely to start a new shift and work for longer relative to those who have recently worked less, controlling for all other covariates. We refer to this phenomenon 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 robustness tests where we include the interaction terms between drivers' work experience and each of the key variables, we find that both income targeting and inertia are stronger among newer drivers (e.g., with fewer than 90 working days). A similar impact of experience is documented by Sheldon (2016). 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 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 Appendix D for further discussion on this topic. 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, including the following: allowing for non-linear 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 mixed-effects survival analysis. 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.

1.6. 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 Appendix E, we conduct a policy analysis to demonstrate how our insights can help policymakers evaluate the impact of a regulation.

1.6.1. 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 re-allocate 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.

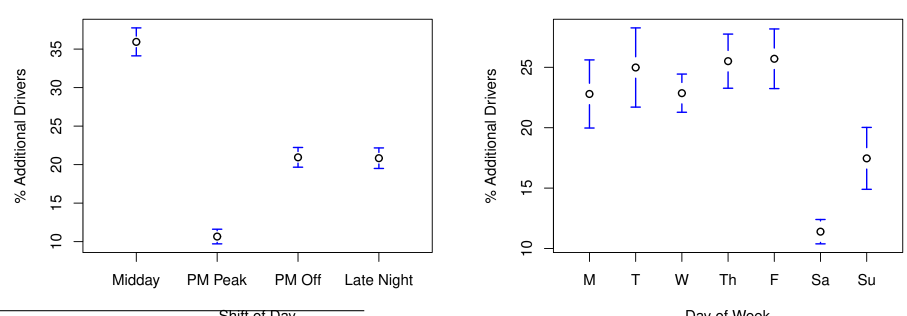
In our context, drivers always receive a guaranteed base pay when they work and sometimes they receive promotions on top of the base rate. We assume that the budget for promotions is separate from the budget for base rates. As not every driver who receives a promotion would choose to work, we compute 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 $\bar{D}$. We then compute the inverse c.d.f. evaluated at $\bar{D}$: $\tilde{D} = \Phi^{-1}(\bar{D})$, that is, $\tilde{D}$ represents the argument of $\Phi(\cdot)$ in the right hand-side of Equation (1.6). In other words, $\tilde{D}$ corresponds to the combination of drivers' attributes that will induce a probability of working equal to $\bar{D}$. 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.

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.

Figure 5: Number of additional drivers using our allocation strategy given a fixed budget

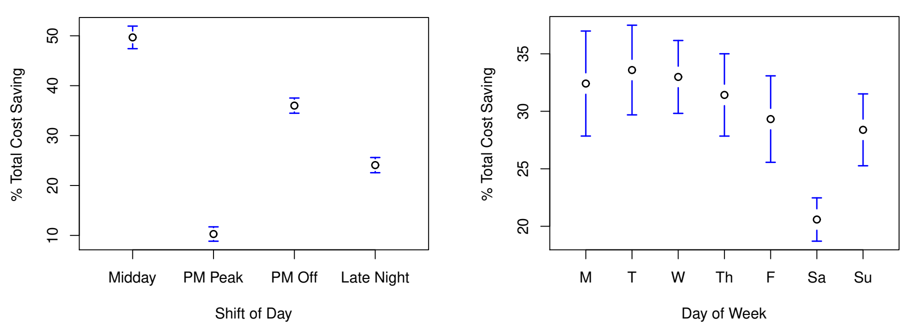


[4]One potential concern is fairness. Future research can include additional constraints such as the minimum fraction of drivers receiving a promotion and the maximum number of different promotion levels.

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.

Figure 6: Simulated cost savings while maintaining the same service capacity



1.6.2. 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:

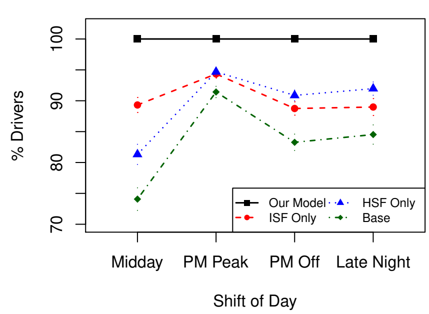
(a) *ISF Only*: The firm assumes that work decisions are influenced by ISF but not HSF.

(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) above 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 §1.6.1, 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 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: Impact of ignoring behavioral factors on the expected number of active drivers



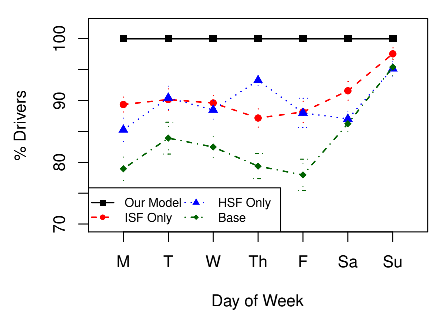


Figure 7 shows that ignoring behavioral factors can lead to a significant loss in the number of active drivers. Specifically, the Base model leads to an average loss of 16.70% in the expected number of active drivers relative to our model, with a standard deviation of 13.06%. The ISF Only (HSF Only)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

1.7. 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. While 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.

One of the important phenomena that emerge from this paper is the existence of inertia among drivers. While 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 delivery to tutoring services. Therefore, there is a lesson to be learned about the fundamental impact of such policies. Amidst 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 confirm insights, testing the income targeting and inertia effects could be of interest. 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. 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 is also an interesting extension. 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 platforms to efficiently motivate and strengthen their relationships with their flexible workforce.

A. Additional Details of the Main Results

Figures A1 and A2 provide additional details of the main results from our two-stage model of drivers' decisions across shifts and across days, respectively. For each of the key variables, we provide the estimated coefficient and the standard error in parenthesis. Within each model, we also report the estimated coefficient and the standard error for IMR and two R^2 values, total R^2 (top) and within R^2 (bottom, italicized). We acknowledge that a few of the IMR estimates are not statistically significant, suggesting that the selection bias is weak in some cases. However, our insights regarding the impact of financial incentives, cumulative income, and cumulative work hours on the decisions of both stages are consistent across different model specifications and selection approaches (e.g., two-part model and Dahl's correction).

Figure A1: Estimates of our two-stage model of drivers' shift-level decisions across different shifts

	Choice (Work or not)				Level (How long)					
SUV	Offer	ISF	HSF	N	Earn	ISF	HSF	IMR	R^2	N
Midday	**0.0024 (0.0006)**	**-0.0173 (0.0036)**	**2.9044 (0.1632)**	124,769	0.001 (0.001)	**-0.008 (0.002)**	**1.826 (0.070)**	**0.271 (0.029)**	0.552 *0.239*	45,329
PM-Peak	**0.0082 (0.0016)**	**-0.0022 (0.0002)**	**0.5102 (0.0082)**	131,910	**0.023 (0.005)**	**-0.0004 (0.0001)**	**0.316 (0.009)**	**0.627 (0.043)**	0.244 *0.092*	39,592
PM-OPeak	**0.0018 (0.0008)**	**-0.0024 (0.0001)**	**0.3436 (0.0048)**	130,651	**0.003 (0.001)**	**-0.0001 (0.00003)**	**0.020 (0.002)**	0.009 (0.011)	0.281 *0.029*	26,699
Late Night	**0.0035 (0.0010)**	**-0.0024 (0.0001)**	**0.2817 (0.0047)**	125,382	**0.025 (0.002)**	-0.0002 (0.0001)	**0.022 (0.011)**	-0.088 (0.054)	0.296 *0.027*	17,137
Sedan	Offer	ISF	HSF	N	Earn	ISF	HSF	IMR	R^2	N
Midday	**0.0068 (0.0008)**	**-0.0309 (0.0056)**	**3.2429 (0.1916)**	113,444	0.001 (0.001)	**-0.007 (0.002)**	**1.058 (0.061)**	-0.029 (0.029)	0.580 *0.206*	20,297
PM-Peak	**0.0109 (0.0019)**	**-0.0014 (0.0004)**	**0.4852 (0.0133)**	117,152	**0.020 (0.004)**	**-0.001 (0.0002)**	**0.116 (0.009)**	**-0.120 (0.034)**	0.273 *0.014*	19,613
PM-OPeak	**0.0031 (0.0010)**	**-0.0028 (0.0003)**	**0.4133 (0.0090)**	124,611	**0.003 (0.0005)**	**-0.0002 (0.00004)**	**0.005 (0.002)**	**-0.098 (0.007)**	0.252 *0.029*	17,025
Late Night	0.0018 (0.0014)	**-0.0021 (0.0002)**	**0.3356 (0.0082)**	124,280	**0.036 (0.004)**	**-0.001 (0.0002)**	**0.063 (0.011)**	**-0.378 (0.048)**	0.304 *0.026*	15,623

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

Figures A3 provides the effect sizes for an average driver at the shift and day levels, respectively, under one of the following conditions: (i) a \$10 increase in hourly offer or earning rate, (ii) a \$10 increase in ISF, and (iii) an additional hour to HSF.

Figure A2: Estimates of our two-stage model of drivers' day-level decisions across different days

	Choice (Work or not)				Level (How long)					
SUV	Offer	ISF	HSF	N	Earn	ISF	HSF	IMR	R^2	N
Tuesday	0.0039 (0.0021)	0.0006 (0.0003)	**0.0581 (0.0137)**	28,883	-0.003 (0.010)	-0.001 (0.001)	0.027 (0.029)	**-1.711 (0.184)**	0.422 *0.037*	9,482
Wednesday	0.0036 (0.0020)	**0.0005 (0.0002)**	**0.0461 (0.0087)**	21,965	-0.001 (0.008)	-0.0003 (0.0005)	0.028 (0.021)	**-1.274 (0.192)**	0.422 *0.040*	10,120
Thursday	**0.0087 (0.0019)**	**0.0005 (0.0001)**	**0.0358 (0.0061)**	29,233	-0.006 (0.008)	-0.0004 (0.0003)	**0.042 (0.014)**	**-0.973 (0.217)**	0.412 *0.046*	9,894
Friday	**0.0069 (0.0019)**	0.00001 (0.0001)	**0.0506 (0.0046)**	20,294	0.013 (0.008)	-0.0004 (0.0002)	**0.055 (0.012)**	0.007 (0.229)	0.436 *0.031*	9,283
Saturday	**-0.0246 (0.0036)**	-0.0002 (0.0001)	**0.0292 (0.0038)**	15,788	-0.002 (0.030)	0.0001 (0.0003)	-0.013 (0.017)	**-2.149 (0.640)**	0.398 *0.045*	4,372
Sunday	**-0.0216 (0.0034)**	**-0.0006 (0.0001)**	**0.0504 (0.0040)**	13,025	**0.049 (0.024)**	0.00005 (0.0004)	-0.032 (0.021)	**-3.102 (0.580)**	0.390 *0.040*	3,240
Sedan	Offer	ISF	HSF	N	Earn	ISF	HSF	IMR	R^2	N
Tuesday	**0.0216 (0.0028)**	0.0008 (0.0007)	**0.0766 (0.0221)**	21,283	**-0.040 (0.015)**	-0.002 (0.002)	**0.070 (0.035)**	**-0.940 (0.141)**	0.564 *0.097*	4,681
Wednesday	**0.0128 (0.0027)**	**0.0016 (0.0004)**	**0.0435 (0.0142)**	23,280	0.015 (0.012)	**-0.002 (0.001)**	**0.122 (0.023)**	**-0.657 (0.150)**	0.567 *0.114*	5,278
Thursday	**0.0115 (0.0026)**	**0.0010 (0.0003)**	**0.0351 (0.0095)**	19,982	-0.002 (0.011)	-0.00004 (0.0005)	**0.052 (0.016)**	-0.254 (0.164)	0.542 *0.100*	5,081
Friday	**0.0173 (0.0024)**	**0.0004 (0.0002)**	**0.0375 (0.0068)**	18,418	-0.009 (0.011)	-0.00002 (0.0004)	**0.026 (0.013)**	-0.321 (0.209)	0.533 *0.067*	4,666
Saturday	0.0035 (0.0049)	-0.0003 (0.0002)	**0.0502 (0.0062)**	15,762	-0.006 (0.028)	-0.0002 (0.0004)	**0.038 (0.014)**	-0.066 (0.311)	0.514 *0.067*	3,817
Sunday	-0.0081 (0.0046)	**-0.0007 (0.0002)**	**0.0626 (0.0063)**	12,602	**0.058 (0.022)**	**-0.001 (0.0004)**	**0.062 (0.015)**	-0.317 (0.342)	0.560 *0.101*	3,065

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

Figure A3: Effect sizes of changes in hourly financial offer, ISF, and HSF on drivers' shift-level decisions

	Change in P(Work) (percentage points)					Change in Duration (minutes)				
SUV	Mean	+1% Offer	+1% ISF	+1% HSF	N	Mean	+1% Offer	+1% ISF	+1% HSF	N
Midday	0.343	**0.079**	**-0.152**	**0.541**	124,769	4.987	*0.026*	**-0.239**	**1.216**	45,329
PM-Peak	0.277	**0.078**	**-0.064**	**0.348**	131,910	2.421	**0.726**	**-0.056**	**1.054**	39,592
PM-OPeak	0.182	**0.049**	**-0.051**	**0.178**	130,651	0.731	**0.057**	**-0.027**	**0.095**	26,699
Late Night	0.117	**0.031**	**-0.031**	**0.093**	125,382	1.996	**0.484**	-0.023	**0.077**	17,137
Sedan	Mean	+1% Offer	+1% ISF	+1% HSF	N	Mean	+1% Offer	+1% ISF	+1% HSF	N
Midday	0.137	**0.034**	**-0.034**	**0.117**	113,444	4.186	*0.026*	**-0.114**	**0.549**	20,297
PM-Peak	0.123	**0.045**	**-0.007**	**0.080**	117,152	2.327	**0.311**	**-0.075**	**0.312**	19,613
PM-OPeak	0.099	**0.031**	**-0.015**	**0.068**	124,611	0.803	**0.035**	**-0.020**	**0.021**	17,025
Late Night	0.071	*0.033*	**-0.011**	**0.054**	124,280	2.167	**0.579**	**-0.153**	**0.220**	15,623

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

B. Alternative Empirical Approaches

B.1. Sample Selection Bias Correction

Dahl's correction.

Following Dahl (2002) and Bray et al. (2019), we use the selection probability as a sufficient statistic for the selection bias. Since, in our context, the choice for each driver is only binary: to work or not, we do not suffer from the curse of dimensionality. Revisiting our level equation (Equation (1.7)),

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

we can substitute IMR (λ) with all basis functions of a B-spline by using the quantiles of work probabilities for all drivers, $\mathbf{P_{work}} = [P(Drive_{i,t} = 1|\mathbf{X_{i,t}}), \forall i]$ as interior knots. Let $\mathfrak{B}(\mathbf{P_{work}}, j)$ be the j^{th} basis function of a degree n B-spline with the quantiles of $\mathbf{P_{work}}$ as m interior knots. Also, we define $\eta_{i,t} = u_{i,t} - \sum_{j=0}^{m+n}\gamma_j\mathfrak{B}(\mathbf{P_{work}}, j)$ to maintain the orthogonality of the error term and the expected hours worked. Thus, our level equation under this approach becomes:

$$f(Hour_{i,t}) = \beta_{0,i} + \beta_{\tilde{w}}\tilde{w}_{i,t} + \beta_{ISF}ISF_{i,t} + \beta_{HSF}HSF_{i,t} + \beta\mathbf{Z_{i,t}} + \sum_{j=0}^{m+n}\gamma_j\mathfrak{B}(\mathbf{P_{work}}, j) + \eta_{i,t}. \tag{B1}$$

In Figure B4, we present the estimates for the level equation when choosing $m = n = 3$. Our results remain consistent under both approaches for sample selection correction. Note that, for all but sedan drivers' decisions on Friday and Saturday, the selection variables are significant at $p = 0.05$, hence confirming that there exists a selection bias in the decision to work.

Figure B4: Estimates for the level equation using Dahl's correction

	SUV Drivers			Sedan Drivers		
Shift	Earning	ISF	HSF	Earning	ISF	HSF
Midday	**+**	+	**+**	**+**	+	**+**
PM-Peak	+	**-**	**+**	**+**	**-**	**+**
PM-OPeak	**+**	**-**	**+**	**+**	**-**	**+**
Late Night	**+**	**-**	**+**	**+**	**-**	**+**
Day	Earning	ISF	HSF	Earning	ISF	HSF
Tuesday	-	-	+	-	-	+
Wednesday	+	-	+	**+**	**-**	**+**
Thursday	-	-	+	+	+	+
Friday	-	**-**	+	-	+	+
Saturday	**-**	+	**-**	-	+	**-**
Sunday	-	**+**	**-**	+	**-**	**+**

Note: Green background with bolded "+": significantly positive, yellow with bolded "-": significantly negative, white with italicized sign: non-significant. All at $p = 0.05$.

B.2. Instrumental Variables

Co-skippers IV.

This IV follows a similar idea to our main IV, but instead of matching drivers based on their past work decisions at a specific time in the past, we now match drivers based on 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 been inactive. We call the drivers of a different vehicle type who belong to the same group *co-skippers*. This IV satisfies the *relevance condition:* Since both the focal driver and their co-skippers have been inactive for approximately the same time, their incentives should be highly correlated. From the first stage of our IV estimation, the estimate for the instrument is consistently signifiant and F-statistics across all models except one are larger than the conventional threshold of 10. This IV also satisfies the *exclusion restriction:* Current incentives for co-skippers should not directly influence the focal driver's work decision because (i) they drive different vehicle types and (ii) the focal driver does not have access to co-skippers' incentives information.

The estimates from shift- and day-level analyses are consistent with our main results. Figure B5 presents the signs and statistical significance (at $p = 0.05$) of the estimates across shifts and days. However, these models are outperformed by our main model both in terms of in-sample and out-of-sample accuracy.

Figure B5: Estimates across shifts and days using the co-skippers IV

	Choice (Work or not)				**Level** (How long)			
SUV	IV-F	Offer	ISF	HSF	IV-F	Earning	ISF	HSF
Midday	433.1	+	+	+	266.8	+	+	+
PM-Peak	289.5	-	-	+	58.7	+	-	+
PM-OPeak	260.2	+	-	+	45.1	+	-	+
Late Night	329.9	+	-	+	36.4	+	-	+
Sedan	IV-F	Offer	ISF	HSF	IV-F	Earning	ISF	HSF
Midday	229.2	+	+	+	104.4	+	+	+
PM-Peak	231.3	+	-	+	31.8	+	-	+
PM-OPeak	255.9	+	-	+	24.3	+	-	+
Late Night	270.0	+	-	+	30.9	+	-	+

	Choice (Work or not)				**Level** (How long)			
SUV	IV-F	Offer	ISF	HSF	IV-F	Earning	ISF	HSF
Tuesday	37.9	-	+	+	21.1	-	-	+
Wednesday	41.3	+	+	+	25.5	-	-	+
Thursday	67.9	+	-	+	43.9	-	+	+
Friday	67.5	+	+	+	41.9	-	-	+
Saturday	89.1	+	+	+	19.4	-	+	-
Sunday	82.1	-	-	+	16.0	+	+	-
Sedan	IV-F	Offer	ISF	HSF	IV-F	Earning	ISF	HSF
Tuesday	25.3	+	+	+	8.8	-	-	+
Wednesday	24.9	+	+	+	10.6	+	-	+
Thursday	43.9	+	+	+	15.8	+	-	+
Friday	39.9	+	+	+	13.6	-	+	-
Saturday	58.3	+	+	+	12.8	-	+	-
Sunday	48.7	-	-	+	12.5	+	-	+

Note: Green background with bolded "+": significantly positive, yellow with bolded "-": significantly negative, white with italicized sign: non-significant. All at $p = 0.05$.

Hausman-type IV.

Inspired by previous studies such as Sheldon (2016), we use the average hourly offer rate received by all other registered drivers during the same shift on the same day as an instrument for the offer rate. Similarly, we use the average hourly earning rate earned by all other active drivers during the same shift on the same day as an instrument for the hourly earning rate. These instruments can be thought of as a mutual offer or earning rate for eligible drivers in New York City at a particular time. In addition, the incentives offered to other drivers should not directly influence the focal driver's decision to work. Controlling for weather and market conditions using the TLC data, we rule out potential confounders that affect both the variation in incentives and in labor decisions. Recall that unlike other ride-hailing platforms, drivers on our platform do not compete with other drivers for promotions as both the base and promotional rates are decided and announced ahead of time. Moreover, promotions are not offered as a way to relocate drivers to high-demand areas (see §1.3.3 for more details). Thus, it suggests that this IV satisfies the exclusion restriction. The results we obtained using this IV are qualitatively similar as illustrated in Figure B6. While this type of IV appears to be valid for the choice equation, low F-statistics suggest that it is a relatively weaker IV relative to both the co-workers and co-skippers IVs.

Figure B6: Estimates across shifts and days using Hausman-type IV

	Choice (Work or not)				**Level** (How long)			
SUV	IV-F	Offer	ISF	HSF	IV-F	Earning	ISF	HSF
Midday	365.4	+	+	+	183.7	+	+	+
PM-Peak	318.8	-	-	+	56.9	+	-	+
PM-OPeak	301.7	+	-	+	49.1	+	-	+
Late Night	371.7	+	-	+	38.1	+	-	+
Sedan	IV-F	Offer	ISF	HSF	IV-F	Earning	ISF	HSF
Midday	223.2	+	+	+	81.1	+	+	+
PM-Peak	255.6	-	-	+	32.9	+	-	+
PM-OPeak	274.9	+	-	+	27.3	+	-	+
Late Night	278.6	+	-	+	29.8	+	-	+

	Choice (Work or not)				**Level** (How long)			
SUV	IV-F	Offer	ISF	HSF	IV-F	Earning	ISF	HSF
Tuesday	41.4	+	+	+	18.1	-	-	+
Wednesday	52.2	+	+	+	24.5	+	-	+
Thursday	61.6	+	-	+	33.2	-	+	+
Friday	62.3	+	+	+	29.8	-	-	+
Saturday	83.1	+	+	+	15.2	-	+	-
Sunday	70.9	+	-	+	11.5	+	+	-
Sedan	IV-F	Offer	ISF	HSF	IV-F	Earning	ISF	HSF
Tuesday	30.3	+	+	+	8.0	-	-	+
Wednesday	35.3	+	+	+	10.3	+	-	+
Thursday	42.6	+	+	+	12.9	+	+	+
Friday	39.5	+	+	+	10.6	-	+	-
Saturday	44.4	+	-	+	9.8	-	+	-
Sunday	40.7	+	-	+	7.9	+	-	+

Note: Green with "+": significantly positive, yellow with "-": significantly negative, white: non-significant at $p = 0.05$.

B.3. Addressing the Multicollinearity Concern

Correlations between ISF and HSF in our data range between 0.446 and 0.929, depending on the time of the day and the vehicle type. While these correlations appear to be on a high side, we gain sufficient statistical power by leveraging our large sample size. Based on Mason and Perreault Jr (1991), our levels of collinearity are between Levels II and III. Given that our R^2 is between 0.25 and 0.5, the minimum sample size of 300 is required. In our case, this requirement is readily satisfied since we have over 100,000 observations for each vehicle type and shift.

Nevertheless, we consider alternative model specifications that still allow us to investigate both the impact of ISF and HSF on the labor decisions. For conciseness, we present one major approach below. The insights remain valid in all specifications.

Localized hazard regressions.

Motivated by Thakral and Tô (2019), we estimate additional models when controlling for drivers who either had the same amount of accumulated earnings or the same amount of time worked so far. Such a specification allows for a flexible, driver-specific hazard of stopping and a time-dependent relationship between each of the covariates and the stopping probability. After driving t trips and accumulating y_{int} from working a total of h_{int} hours, driver i decides to end shift n when the cost of additional effort exceeds the expected continuation value. The variables y_{int} and h_{int} represent income so far (ISF) and hours so far (HSF) in our setting. We let d_{int} be the decision to stop working after trip t in shift n. Thakral and Tô (2019) models the probability that driver i ends shift n at trip t by

$$\mathbb{P}(d_{int} = 1) = f(h_{int}) + \beta(h_{int})y_{int} + X_{int}\gamma(h_{int}) + \mu_i(h_{int}) + \epsilon_{int},$$

where $f(\cdot)$ represents the baseline hazard and μ absorbs differences in drivers' baseline stopping tendencies. HSF affects the stopping probability through the baseline hazard and the impact of ISF, covariates, and drivers' fixed effects. $\beta(h)$ reflects the effect of an additional dollar of ISF on the probability of ending a shift for a driver after h hours of work ($HSF = h$). Thakral and Tô (2019) employs local linear regressions to estimate the baseline hazard and the time-varying coefficients by solving a separate weighted least squares problem:

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

with weights given by $w(\cdot)$. With uniform weights, this procedure becomes fitting a linear model to a localized subset of data. We consider time windows of different interval: 10, 15, 20, 30, and 60 minutes.

Specifically, we consider the following two models:

(i) *HSF impacts how ISF affects the stopping probability.* This is similar to the model formulated in Thakral and Tô (2019). We model the probability that driver i stops working at time t of day n after earning ISF_{int} and spending HSF_{int} hours working for the day as:

$$\mathbb{P}(d_{int} = 1) = f(HSF_{int}) + \beta^{w}(HSF_{int})w_{int} + \beta^{ISF}(HSF_{int})ISF_{int} \\ + X_{int}\gamma(HSF_{int}) + \mu_i(HSF_{int}) + \epsilon_{int},$$

where w_{int} is the hourly financial incentive offered at time t of day n. We include the hourly incentive to match our main models and reflect the possibility that drivers are less likely to quit if the current offer is appealing. The local regressions are done by controlling for drivers who were still active at the population median of HSF.

(ii) *ISF impacts how HSF affects the stopping probability.* This model is to validate our findings that drivers exhibit inertia, affecting their work decisions. Using the notation from our setting, we model the probability that driver i stops working at time t of day n after earning ISF_{int} and spending HSF_{int} hours working for the day as:

$$\mathbb{P}(d_{int} = 1) = f(ISF_{int}) + \beta^{w}(ISF_{int})w_{int} + \beta^{HSF}(ISF_{int})HSF_{int} \\ + X_{int}\gamma(ISF_{int}) + \mu_i(ISF_{int}) + \epsilon_{int}.$$

The local regressions are done by controlling for drivers who were still active when earning cumulative income of the population median of ISF.

Results for Model (i): Impact of ISF. The median number of hours that drivers worked on non-holiday weekdays is 6.72 for SUV drivers and 6.58 for sedan drivers. Table B1 presents the estimates for the local probit models of the decision to quit within 10, 15, 30, or 60 minutes after reaching the population median HSF. The results confirm that financial incentives decrease the quitting probability, while cumulative earnings tend to increase the quitting probability. Under the assumption that cumulative hours worked (HSF) only affect the quitting probability through the impact of offers and ISF, we confirm that income targeting exists while drivers appear to have a positive income elasticity.

Table B1: Estimates of local probit models of quitting decision controlling for cumulative work hours (HSF)

Quit within	SUV		Sedan	
	Offer	ISF	Offer	ISF
10 mins	-0.0174	0.0004	-0.0340	0.0025
15 mins	-0.0199^{*}	0.0014	-0.0365^{*}	0.0040^{*}
30 mins	-0.0204^{**}	0.0023^{*}	-0.0321^{**}	0.0039^{**}
1 hour	-0.0047	0.0011	-0.0165	0.0016
Note:		*p<0.05; **p<0.01; ***p<0.001		

Results for Model (ii): Impact of HSF. We perform a similar analysis where we assume that the impact of ISF is only through the varying impact of HSF. The median cumulative earnings drivers made on non-holiday weekdays are \$219.73 for SUV drivers and \$199.01 for sedan drivers. Table B2 shows that significant inertia is observed among SUV and sedan drivers when the time window of quitting decision is between 10 and 30 minutes. We also find that the hourly financial offer consistently decreases the stopping probability except for large SUV drivers where the effect is the opposite.

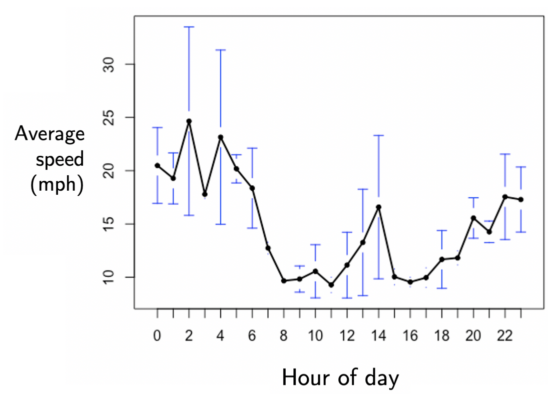
Table B2: Estimates of local probit models of quitting decision controlling for cumulative earnings (ISF)

Quit within	SUV		Sedan	
	Offer	HSF	Offer	HSF
10 mins	-0.18	-0.0652	-0.0047	0.0349
15 mins	-0.0252^{***}	-0.1003^{***}	-0.0091	0.0019
30 mins	-0.0186^{***}	-0.0718^{***}	-0.021^{***}	-0.1103^{***}
1 hour	-0.0202^{**}	-0.0235	-0.0182^{***}	0.0228
Note:		*p<0.05; **p<0.01; ***p<0.001		

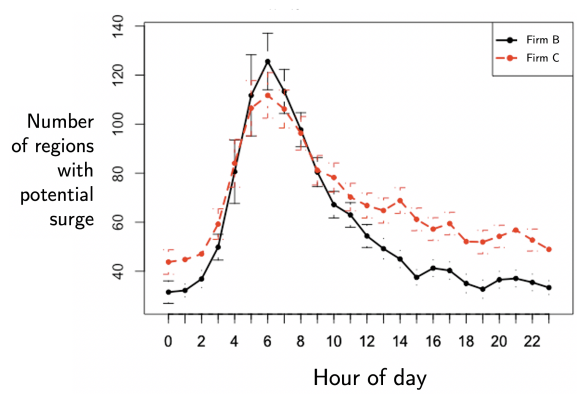
B.4. Alternative Construction of ISF and HSF

We first argue that our assumption that the progress toward a daily income or time goal is reset at midnight is reasonable. 91.07% of drivers' working days observed in our data do not overlap with midnight (e.g., they did not work overnight). Furthermore, 99.93% started working between 5am and 11pm. Therefore, we believe that drivers consider a new calendar day as a new progress. However, it is plausible that drivers do not reset their weekly goals every Monday. As a robustness study, we relax the assumptions that the weekly targets are reset every Monday. Instead, drivers might reset the across-day goals only when they start working after being inactive for some time. In this direction, we analyzed the duration of inactivity between any consecutive working days. Among 7,800 drivers who worked at least two days in our dataset, the average number of inactive days between two working days is 2.21. 15% drivers worked everyday on average and 53.30% did not take more than 2 days break. We re-estimated our models by allowing the targets to be reset every time the drivers did not work for at least two days. Allowing the weekly targets to be reset after taking time off from work, our original insights remain qualitatively consistent.

Figure B7: Additional competition metrics by hour of day for non-holiday weekdays



(a) *Speed*: Average speed in miles per hour



(b) *AggSurge*: Number of NYC regions with potential surge pricing on competing platforms

C. Competition Among Ride-hailing Platforms

In §1.4.2, we discussed four different metrics to control for unobserved demand for ride-hailing services and competition effects. Our main results presented in §3.5 include all observations from October 2016 to September 2017, the weather information, and the aggregated number of trips on competing platforms ($NumFHV$) as controls for market conditions. For observations between July and September 2017, we conduct an additional analysis to further include $Speed$ and $AggSurge$ as covariates. Figures B7a and B7b illustrate the variations in $Speed$ and $AggSurge$ by hour of day for non-holiday weekdays, respectively.

These new results are qualitatively consistent with our main results. Tables C3 and C4 display the estimates for the first-stage estimation of whether or not to work for each shift. We observe a generally positive income elasticity, income targeting behavior, and inertia throughout all the shifts. Speed appears to have a negative impact on the decision to work in general, suggesting that drivers are less likely to work for the focal platform when there is less traffic. The aggregated surge has also a negative impact on the decision to work. This is to be expected: given that the financial incentive for the focal platform is fixed and known, drivers are less likely to work when the outside option is more appealing.

Table C3: Estimates for the shift-level first-stage estimation for sedan drivers during Summer 2017

Sedan	Offer	ISF	HSF	Speed	AggSurge
Mid-day	0.0075***	-0.0354**	3.6385***	-0.0298	-2.9551***
PM peak	-0.0209***	-0.0016*	0.4743***	-0.0536**	-3.9532***
PM off-peak	0.0136***	-0.0034***	0.413***	0.0132	-1.1326**
Late night	0.01079**	-0.004***	0.38036***	-0.07055***	-0.51665
Note:					*p<0.05; **p<0.01; ***p<0.001

Table C4: Estimates for the shift-level first-stage estimation for SUV drivers during Summer 2017

SUV	Offer	ISF	HSF	Speed	AggSurge
Mid-day	0.0035**	-0.0535***	4.3936***	0.0007	-2.5716***
PM peak	-0.0433***	-0.0024***	0.5249***	-0.0563***	-3.6690***
PM off-peak	0.0028	-0.0024***	0.3414***	-0.0121	-0.2124
Late night	0.0085***	-0.0023***	0.2945***	-0.0785***	0.0920
Note:					*p<0.05; **p<0.01; ***p<0.001

The results for the second stage are relatively consistent as well (see Tables C5 and C6). Higher hourly earnings appear to be associated with a longer work duration for most shifts. Income targeting behavior becomes less significant. Inertia is stronger earlier on in the day. Finally, we observe that, conditional on driving for the shift, drivers are less influenced by the traffic conditions or by the potential surge pricing from other platforms.

For the day-level analysis, we find that, in the first-stage estimation, positive income elasticity and income targeting behavior became less apparent. Sedan drivers responded positively to the hourly offer from Tuesday to Thursday, whereas SUV drivers did not. The effect of

Table C5: Estimates for the shift-level second-stage estimation for sedan drivers during Summer 2017

Sedan	Earnings	ISF	HSF	Speed	AggSurge	IMR
Mid-day	0.008	-0.019***	1.604***	-0.039	0.0003	***
PM peak	0.025*	-0.001	0.084***	0.012	0.029	***
PM off-peak	0.003	-0.003***	0.006	-0.0001	0.147	***
Late night	0.03***	0.001	-0.071**	0.019	0.11*	***
Note:	*p<0.05; **p<0.01; ***p<0.001					

Table C6: Estimates for the shift-level second-stage estimation for SUV drivers during Summer 2017

SUV	Earnings	ISF	HSF	Speed	AggSurge	IMR
Mid-day	-0.001	-0.008	1.819***	-0.034	-1.008	***
PM peak	0.062***	-0.0002	0.245***	-0.053***	-0.421	***
PM off-peak	0.004***	-0.0002*	0.033***	-0.002	-0.027	***
Late night	0.022***	0.0001	0.021	-0.006	0.843	
Note:	*p<0.05; **p<0.01; ***p<0.001					

cumulative earnings is generally insignificant, except a sign of income targeting at the end of the week. However, inertia is still significant and apparent for most days, Thursday through Sunday for sedan driver, and Wednesday through Sunday for SUV drivers. Lastly, for the second-stage estimation, we find no significant estimates for our key variables. This is in line with our original results, which led us to conclude that the decision on the work duration for the day was not determined at the beginning of the day.

D. Psychological Explanations for Our Main Results

Our main results suggest that workers on our focal platform exhibit different behaviors regarding cumulative earnings and recent work duration. We believe such different behaviors stem from the fact that people perceive the value of time and money differently. Contrary to a common saying that time is money, empirical research from psychology shows that decisions about time follow different rules than decisions about money. For example, Leclerc et al. (1995) finds that people are more averse to uncertainty with time as contrasted with money. In other words, people are risk averse with respect to decisions in the domain of time loss despite being risk-seeking with respect to decisions involving monetary loss. The authors concluded that because time is less substitutable than money, being certain is more important for decisions about time, and people are more averse when there is uncertainty about the allocation of time. Soman (2001) shows that people do not mentally account for their time in the same way as they account for money as the former is more difficult, while Okada and Hoch (2004) demonstrates that people spend time in a systematically different way from spending money because the value of time is of greater ambiguity. The distinction of attitude toward time and money applies to work motivation and decisions as well. Workers who can adjust their own work schedules are found to be influenced by internal reference targets. Depending on the context, workers may form only a target for income (Camerer et al. 1997), a target for time (Farber 2015), both in the same direction (Crawford and Meng 2011), or both in the opposite direction as observed in our work. DeVoe and Pfeffer (2007) shows that organizational practices such as how firms pay their employees may influence employees' psychological evaluation of time and the tradeoffs they make between time and money.

Our key insight suggests that gig economy workers may exhibit inertia at work. In our context, inertia refers to the positive correlation between the recent work duration and the decision to start a new work shift. We have identified the following three potential explanations of inertia from the fields of psychology, organizational behavior, and management.

(i) First, inertia could be linked to the concept of experience of flow from positive psychology. A flow state is the mental state in which a person performing an activity is fully immersed in a feeling of energized focus, full involvement, and enjoyment in the process of the activity (Csikszentmihalyi and Csikszentmihalyi 1992). The complete absorption into the activity affects how the person perceives the sense of time, leading to a continuation of performing the task even though the marginal benefit is negligible. Flow theory postulates key conditions required to achieve a flow state. These conditions include clear goals and task structure, clear and immediate performance feedback, a balance between the challenges of the task and one's own skills, one's feeling of control, and one's intrinsic motivation. Gig economy workers are likely to meet these conditions since gig tasks typically have a known set of goals and structure, feedback (e.g., from customers) and compensation are provided frequently, and workers are generally skilled at the particular tasks and have some control over their decisions (e.g., work schedule). Csikszentmihalyi and LeFevre (1989) suggests that flow can be experienced in both work and leisure settings, but more dominantly in the former. Among different leisure activities, the authors find that driving is the most common task that generates the flow experience. This finding fits well with our analysis of ride-hailing drivers. Therefore, it is possible that drivers on our platform are more likely to work if they recently worked for a longer duration because they are more likely to experience the flow state.

(ii) Second, inertia may reflect work addiction caused by stochastic rewards. Applying insights from neuroscience research that stochastic rewards could act as a motivator, Corgnet et al. (2020) conducts a series of behavioral experiments to investigate the relationship between stochastic rewards and workers' likelihood to quit working on effortful tasks. The authors found that participants who were offered a stochastic rate of compensation stayed working for a longer period than those offered a deterministic rate. The persistence on the tasks is linked to stress generated by the uncertainty. In a gig economy setting, compensation to workers is typically determined in response to real-time market conditions (e.g., demand) and depends on the specific task and workers' performance. Work addiction among gig workers has been documented and attributed to the rate of compensation (Kruzman 2017). For our focal platform, financial incentives are decided and communicated to drivers ahead of time, but drivers' opportunity costs (e.g., incentives from competing platforms) are not deterministic. Therefore, it is possible that inertia is related to workaholism driven by uncertain rewards.

(iii) Third, inertia, as the absence of fatigue, could be associated with gig workers' flexibility in deciding work schedule. Watanabe and Yamauchi (2016) shows that when workers voluntarily opted to work for a longer period, there is a positive effect on their work-life balance due to the enjoyment of the work itself or increased rewards. Having control over work duration and being compensated for the work are found to be important for workers' satisfaction. Similarly, workers who voluntarily chose to work overtime did not feel more fatigued but instead felt satisfied as long as they chose their own schedule (Beckers et al. 2008). Although the concept of overtime work can only be applied loosely to gig workers since they have full control of their entire schedule, these findings highlight the potential beneficial impact of the flexibility to choose one's own work schedule: reduced fatigued and increased satisfaction. A study on technical contractors whose schedule were not decided by the organization shows that, despite having full control over their work schedule and perceiving the privileged flexibility, these contractors chose to work long hours and appeared to follow a less flexible schedule (Evans et al. 2004). They considered leisure time as a period of loss without pay and hence they sought to minimize time away from work. Using the British Household Panel Survey, DeVoe et al. (2010) observes that individuals who received hourly wage are more willing to trade their leisure time to work and earn more money than those receiving a salary pay. Putting these findings together, we conclude that in our setting where workers can freely choose their own work schedule and receive a hourly pay, they are more likely to work for longer, become more satisfied with long work hours, and feel less fatigue.

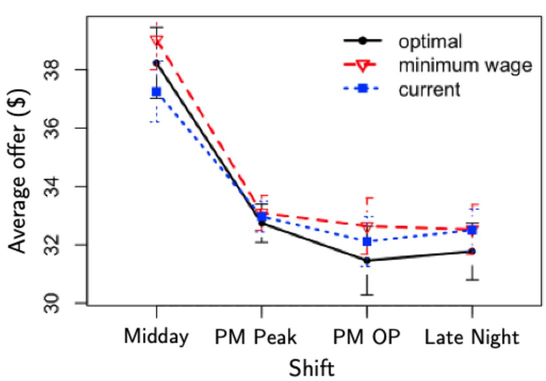
E. Policy Analysis: NYC's Driver Income Rules

Here, we take the perspective of a policymaker and leverage our insights to evaluate the impact of regulations on the welfare of gig workers. In December 2018, the TLC passed *Driver Income Rules* to protect driver earnings, requiring ride-hailing platforms to compensate drivers by a minimum amount for each trip at the rate equivalent to $27.86 per hour. Since there were no such rules during the timeframe of our data, we can only perform a counterfactual analysis to quantify the impact of this new regulation on the workers' welfare, particularly on their earnings.

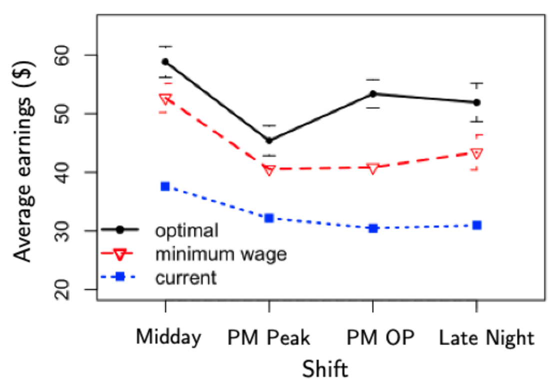
We compare three different policies. First, our *optimal* policy is the targeted incentive allocation policy introduced in §1.6.1, which optimizes incentives based on drivers' predicted probability to work. Second, a *minimum wage* policy adds a constraint to the optimal policy such that every driver must be guaranteed a minimum hourly offer of \$27.86. Finally, we use the observed incentives in the data as a benchmark or *current* practice. Outcomes of interest are the average hourly offer across all drivers and the average hourly earnings across drivers who are predicted to work. The counterfactuals are performed using the data between January and September 2017 in the same fashion as in §1.6.1.

Figure E8a shows that the minimum wage policy slightly increases the average hourly offer among drivers compared to the optimal policy and to current practice, but the differences are not statistically significant. However, these policies lead to significantly different average hourly earnings among drivers predicted to work. Figure E8b suggests that, compared to the current practice, the minimum wage policy significantly improves the average hourly earnings. However, drivers could have earned 10 to 23% more per hour if the incentives were optimally allocated by following the optimal policy without the minimum wage constraint.

Figure E8: Average hourly offer and earnings across three policies



(a) Average hourly offer across all drivers



(b) Average hourly earnings across drivers who worked

The minimum wage policy appears to be beneficial to the workers compared to the platform's current practice. However, as firms are becoming more data-driven and potentially adopting more sophisticated incentive policies (such as our proposed optimal policy), the current minimum wage rule may no longer improve the welfare of the workers. In this case, if the focal platform implements the optimal policy, the regulation decreases workers' pay on average. This also highlights the importance of understanding how gig workers make decisions. The TLC does have detailed information regarding trips operated by ride-hailing drivers but may not have access to how platforms allocate incentives or how drivers decide their flexible schedules. Without such knowledge, policymakers are prone to regulations that could be suboptimal.

CHAPTER 2 : The Structural Behavioral Model of Gig Economy Workers

2.1. Introduction

With the flexibility in the choice of service, gig workers often exhibit a "multihoming" behavior. The majority of ride-hailing drivers work for more than one platform and many also provide other services such as food delivery. An increase in the number of available options has resulted in increased competition among platforms to win over a limited mutual pool of workers. Such competition has only been further exacerbated in cities like New York and Seattle, which recently passed caps on the number of ride-hailing drivers to solve congestion problems and ensure drivers' welfare (Fitzsimmons 2018, Wilson 2019). How workers respond to platform competition is therefore an important topic to study, but studying multihoming behavior empirically is challenging due to the unobservability of workers' available work options. A handful number of recent papers have offered theoretical predictions about the impact of multihoming behavior on the workers. On a positive side, workers benefit from multihoming as they spend less time idling (Liu et al. 2017, Bryan and Gans 2019). However, in the equilibrium, they end up working more but earning less (Benjaafar et al. 2020). In addition, less is known about the magnitude of multihoming behaviors and how firms can influence workers' decisions regarding *where* to work. Closest to our work, Rosaia (2020) presents a model of competing transportation platforms and estimates profit-maximizing prices and the impact of a merger on idle vehicles and efficiency. In this work, we leverage proprietary data from our ride-hailing industry partner and the publicly available trip record data from New York City's Taxi and Limousine Commission to develop and estimate a structural model of gig workers' sequential dynamic decisions in the presence of alternative work opportunities.

We first develop our structural model of gig workers' labor decisions of when and where to work based on time and their current location. We consider a setting where there are two gig economy firms (e.g., our industry partner and a competing firm) that each worker can choose to work for. From the beginning of the time horizon (e.g., the beginning of each working day), the worker chooses when to start working and which firm to work for. On a particular platform, they perform a task (e.g., driving customers to a destination), observe market information (e.g., demand for service at the current location and time, potential future location and earnings), and decide whether to continue working on the focal firm, switch to the competing firm, or stop working. We assume that each driver has a perceived cost of working for each time unit drawn from a population distribution of costs and that their decision-making follows a single-agent dynamic optimization framework. Combining the proprietary dataset on each driver's first and last trips of each active work session with the publicly available trip record data, we first estimate expected values for working for each firm at different times and locations and then simulate for each day the path of decisions of each driver and the time and location s/he ends the session. The simulation relies on the time- and platform-specific incentive structure and transition matrices calibrated from the data. Then, our estimation procedure is based on a likelihood-free approximate Bayesian computation with a sequential Monte Carlo sampler (Sisson et al. 2007) and a machine learning-based adversarial indirect inference estimation (Gourieroux et al. 1993, Kaji et al. 2020) to obtain posterior distribution of the key parameters that generate the simulated outcomes to be as close as possible to the observed decisions. With our parameter estimates, we then perform counterfactual analyses to demonstrate the effectiveness of different strategies commonly used in practice and offer insights that can help the firm manager their workers for different demand scenarios.

Our results characterize workers' forward-looking behavior and heterogeneous perceived cost of working. We observe that a substantial portion of drivers on the focal platform exhibit a multihoming behavior and that drivers can be clustered into low- and high-cost drivers. We also find that drivers are strategic in their choice of initial service location to ensure high utilization and are prone to multihoming when facing longer idle times. Since drivers are paid regardless of the number of customers on the focal platform, our finding potentially suggests that serving customers provides non-monetary benefit to the drivers, either psychologically or by routing them to other high-demand locations. The natural follow-up question is how the firm can reduce (or induce in some cases) multihoming behavior. We consider two policies inspired by practice: (i) using consecutive work bonus (e.g., earn extra money after three consecutive trips), and (ii) imposing a time delay to quit (e.g., drivers need to request to end the session ahead of time). Our counterfactual analyses show that these policies can be effective depending on the goal of the platform. During peak time when the firm wants to retain workers to stay on their platform, using the consecutive work bonus lowers the switching probability and increases work duration on the focal platform. Providing such incentive potentially creates a goal or a sense of commitment to the workers. On the other hand, during the low demand period when fewer workers are needed, introducing the time delay can nudge workers to leave the focal platform sooner. This policy makes being idle and getting stuck in a low-demand region more salient, thus workers would be encouraged to quit earlier than expected. Our major contributions are in the modeling and estimation of dynamic decisions with temporal and spatial components and dynamic outside options, and the development of an efficient simulation-assisted estimation framework in the presence of analytically or computationally intractable likelihood functions and high-dimensional data.

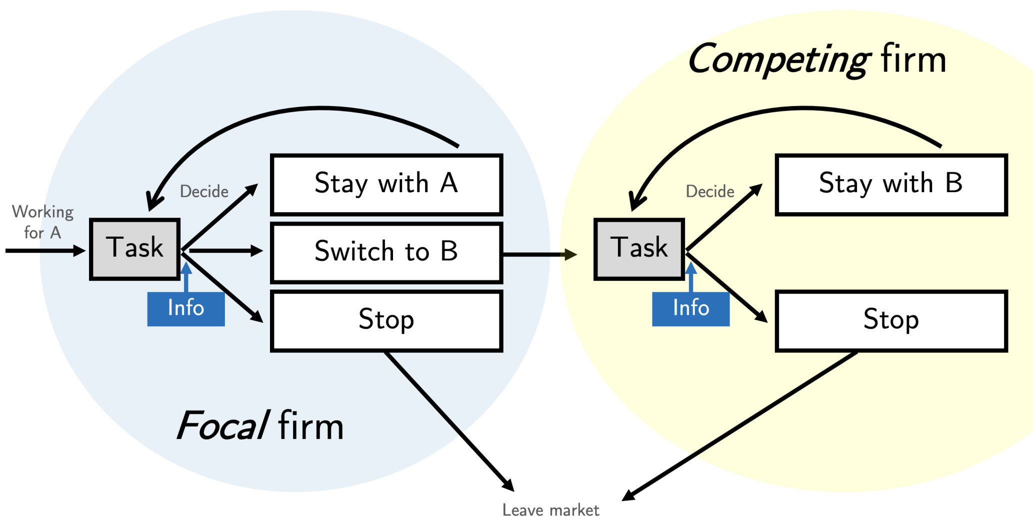
2.2. The Model of Gig Workers' Decisions

We consider a market with two gig platforms offering similar service (e.g., ride-hailing). The two platforms are denoted as the focal firm A and the competitor firm B. This market covers one city with L regions and we consider a discrete time horizon as intervals of 20 minutes from 7am to midnight. For the estimation, we consider the following time blocks: AM peak (7-9am), Late AM (9am-12pm), Midday (12-2pm), Early PM (2-5pm), PM Peak (5-8pm), PM off-peak (8-9pm), Late night (9pm-midnight). We assume that there is no waiting time so the match between the active worker and the customer happens instantly if available.

We focus on drivers who already choose to work for the focal firm A first. After completing each task on A, they observe new information about the market such as demand for various services and make a decision whether to continue on A, switch over to the competitor firm

B, and stop working. If switching over to B, they enter a similar model: after completing each task on B, they observe new information and make a decision whether to stay or quit. We assume that once the drivers switched to B, they will not switch back to A. Figure 9 illustrates the high level overview of our model.

Figure 9: Overview of Our Model



We characterize the drivers by two parameters. All drivers have the same time discounting factor $\beta < 1$ that represents how forward-looking they are. Each driver i has a fixed cost to work C_i per each time interval. $C_i \sim F(\cdot)$ essentially determines the type of drivers they are. Next, we present how we model drivers' decision when working for each of the two platforms.

2.2.1. Working for the Competitor Firm B

A competitor firm B's driver i at location l and time t decides whether to work or not. B drivers can make a decision at the beginning of each time interval if currently not working, not having a passenger in the previous time interval, or just dropping off a passenger.

- f_{lkt} is an average fare of a trip from l to k at t.
- π_{lkt} is a probability that a passenger from l is going to k at t with B.

- R_{lt} is a probability that an B driver would get a ride at l at time t.

- τ_{lk} is an average duration of an B trip starting from l to k.

First, denote the value of quitting to home as V_{ilt}^{BQuit} If choosing to work for B, the driver will be matched with a passenger with a probability R_{lt}. If matched, the driver will go to a destination k with probability π_{lkt} and the trip will take τ_{lk} time units. If not matched, the driver stays at the same location and makes a decision again at the beginning of the next time interval $t+1$.

$$\begin{aligned} V_{ilt}^{BWork} = -C_i + R_{lt} \left(\sum_{k=1}^{L} \pi_{lkt} \beta^{\tau_{lk}} \left(f_{lkt} - (\tau_{lk} - 1)C_i + \max \left(V_{ik(t+\tau_{lk})}^{BWork}, V_{ik(t+\tau_{lk})}^{BQuit} \right) \right) \right) \\ + (1 - R_{lt})(\beta \max \left(V_{il(t+1)}^{BWork}, V_{il(t+1)}^{BQuit} \right)) \end{aligned} \tag{2.1}$$

2.2.2. Working for the Focal Firm A

A focal firm A's drivers i at location l and time t makes a decision whether to (i) work for A, (ii) switch to work for B, or (iii) go home. A drivers can only make decisions after completing a trip or when there is no passenger onboard.

- A drivers get paid an hourly rate w_{it} if they are active.

- π_{lkt}^{V} is a probability that a A driver will be routed from l to k at t.

- R_{lt}^{V} is a probability that a A driver would get a passenger at l at time t.

- τ_{lk}^{V} is an average duration of a A trip starting from l to k.

If choosing to work for A, the driver will be matched with a passenger with probability R_{lt}^{V}. If matched, the driver will go to a destination k with probability π_{lkt}^{V} and the trip will take τ_{lk}^{V} time units. If not matched, the driver stays at the same location and makes a decision again at the beginning of the next time interval $t+1$.

The driver choose option $j \in \{Work, Switch, Quit\}$. The value function is

$$V_{ilt} = \max\left[V_{ilt}^{AWork} + \epsilon_{ilt}^{w}, V_{ilt}^{Switch} + \epsilon_{ilt}^{s}, V_{ilt}^{Quit} + \epsilon_{ilt}^{q}\right], \tag{2.2}$$

where ϵ_{ilt}^{j} is interpreted as a component of utility of a choice j at location l in time interval t which is known to the driver i but not by us. ϵ_{ilt}^{j} is assumed to have a multivariate extreme value distribution.

$$\begin{aligned}
V_{ilt}^{AWork} &= (w_{it}/3) - C_i + R_{lt}^{V}\left(\sum_{k=1}^{L} \pi_{lkt}^{V} \beta^{\tau_{lk}^{V}} \left((\tau_{lk}^{V} - 1)(w_{it}/3 - C_i) + V_{ik(t+\tau_{lk}^{V})}\right)\right) && (2.3)\\
&+ \beta(1 - R_{lt}^{V}) V_{il(t+1)} \\
&= (w_{it}/3) - C_i && (2.4)\\
&+ R_{lt}^{V}\left(\sum_{k=1}^{L} \pi_{lkt}^{V} \beta^{\tau_{lk}^{V}} \left((\tau_{lk}^{V} - 1)(w_{it}/3 - C_i) + \log\left(\sum_{j\in(w,s,q)} \exp\left(V_{ik(t+\tau_{lk}^{V})}^{j}\right)\right)\right)\right) \\
&+ \beta(1 - R_{lt}^{V}) \log\left(\sum_{j\in(w,s,q)} \exp\left(V_{il(t+1)}^{j}\right)\right) && (2.5)\\
V_{ilt}^{Switch} &= V_{ilt}^{BWork} && (2.6)\\
V_{ilt}^{Quit} &= 0 && (2.7)
\end{aligned}$$

2.3. Data: Ride-hailing Trips in New York City

Our main dataset consists of three sources of data. First, we obtain a proprietary data from our industry partner, a U.S. ride-hailing platform, that includes detailed information about the first pick-up and the last drop-off of every consecutive work session by drivers on their platform. A consecutive work session is defined by the time the driver remains online on the platform until she logs off. This data includes the exact timestamps and GPS coordinates of when and where the first and last trips happened for each consecutive work session. The data spans three months, from July 2017 to September 2017, and consists of approximately

140,000 consecutive work sessions across 6,724 drivers. We combine this data with a large comprehensive dataset of driving activities and financial incentives on the same platform. This data includes each driver's vehicle type, experience with the platform, number of hours driven, and financial incentives offered and earned for each time block of the day. Finally, we leverage a trip record data collected by New York City's Taxi and Limousine Commission (TLC). This data consists of date, time, and location of each pick-up and each drop-off and the dispatching base associated with a ride-hailing platform. This public data allows us to capture the real-time market conditions and estimate missing information regarding competition and outside opportunities.

2.4. Estimation Strategy and Implementation

Our context is New York City and we consider $L = 20$ regions within the city: Bronx, Brooklyn, Newark Liberty International Airport, Central Park, Chelsea, Downtown, Governors Island, Gramercy, Harlem, Lower East Side, Lower West Side, Midtown, Morningside Heights, Upper East Side, Upper West Side, Upper Manhattan, John F. Kennedy International Airport, LaGuardia Airport, Queens, and Staten Island. Each ride-hailing driver i has their private fixed cost to work C_i per unit time which is drawn from a truncated Normal distribution $\mathcal{F}$ with support $[\underline{C}, \bar{C}]$, mean μ_C and variance σ_C^2. Every driver has a discount factor $\beta \in [0, 1]$ for future utility. Therefore, we estimate three population parameters: μ_C, σ_C, β.

2.4.1. Pre-Computation

We first discuss our pre-computation steps where we compute several key values from our data prior to estimating for the key population parameters. Based on a 2014 report by the American Automobile Association, the cost per mile is 59.2 cents for an average sedan and 73.6 cents for an SUV. The 2016 average speed of a car in Manhattan is 7.44 mph. The hourly cost is then \$4.40 for a sedan and \$5.48 for an SUV. In our estimation we consider a 20-minute interval as an average ride-hailing trip lasts 20 minutes. μ_C is then expected to be around \$1.47–\$1.83 per interval. We set $\underline{C} = 0$ and $\bar{C} = 5$.

Expected values of working for B. We consider $C_i \in [0,5]$ with an increment of 0.05 and $\beta \in \{0.8, 0.825, 0.85, 0.875, 0.9, 0.925, 0.95, 0.975, 0.98\}$. For each combination of (C_i, β), we compute the value of working for B starting in location l and time block b either on a weekday or a weekend: $V^B_{l,b,weekday}(C_i, \beta)$ or $V^B_{l,t,weekend}(C_i, \beta)$. Recall that for an active B driver, we assume that they choose among two options: continue working or quit. We assume that the value of quitting in zero at any time and any location. We omit the subscript for day of week for simplicity. For each day type, we solve for the values backward from the last (7th) time block in which all drivers are expected to quit at the end of the block (i.e., midnight).

$$V^B_{l,7,midnight}(\cdot) = 0 \quad \forall l = 1, ..., L$$

We assume that the probability of getting a ride R, the transition probability π, the trip duration τ, and the average B fare f are fixed and computed by the averages within each location, time block, and day of week. Then, we compute the value of working for B in the earlier intervals based on the following equation

$$\begin{aligned} V^B_{l,7,t}(C_i, \beta) = R_{l,7} & \left(\sum_{k=1}^{L} \pi_{l,k,7} \beta^{\tau_{l,k,7}} \left(f_{l,k,7} - (\tau_{l,k,7} - 1)C_i + \max\left(V^B_{k,7,(t+\tau_{l,k,7})}(C_i, \beta), 0 \right) \right) \right) \\ & + (1 - R_{l,7}) \left(\beta \max\left(V^B_{l,7,(t+1)}(C_i, \beta), 0 \right) \right) - C_i \end{aligned}$$

until we obtain convergence: $V^B_{l,7,t}(\cdot) = V^B_{l,7,t+1}(\cdot) \quad \forall l$. The $L \times 1$ vector of converged values reflect the stationary expected value of working for B during the focal time block: $\mathbf{V^B_7} = (V^B_{1,7}, ..., V^B_{L,7})$. This vector is then used as a terminal condition for the earlier time block.

$$V^B_{l,6,T}(\cdot) = V^B_{l,7}(\cdot) \quad \forall l = 1, ..., L$$

In the same fashion, for any time block $b < 7$, we compute intermediate value of working for B by:

$$V^B_{l,b,t}(C_i,\beta) = R_{l,b}\left(\sum_{k=1}^{L}\pi_{l,k,b}\beta^{\tau_{l,k,b}}\left(f_{l,k,b} - (\tau_{l,k,b}-1)C_i + \max\left(V^B_{k,b,(t+\tau_{l,k,b})}(C_i,\beta),0\right)\right)\right)$$
$$+ (1-R_{l,b})\left(\beta\max\left(V^B_{l,b,(t+1)}(C_i,\beta),0\right)\right) - C_i$$

The converged values are then the stationary values for the time block to be used as a terminal condition for the next (earlier) time block.

$$V^B_{l,b,T}(\cdot) = V^B_{l,b+1}(\cdot) \quad \forall l = 1,...,L, \quad \forall b = 1,..,6$$

.

Expected values of working for A. For each A driver i, we observe whether they worked on a particular date d, $Work_{i,d}$, and if so, when and where they started (t^s_i, l^s_i) and ended (t^q_i, l^q_i) their work session. Similar to the estimation of B values, we assume that the probability of getting a ride R, the transition probability π, and the trip duration τ, as well as B values V^B, are fixed within each location, time block, and day of week. However, we will estimate the *driver-* and *date-specific* value of working for A for every 20-minute interval starting from the time the driver started working until midnight for each date that the driver worked. In July 2017, there are 11,109 driver-date pairs.

Due to a much larger computational cost, we consider a sparser grid to compute A's value for each driver: 13 $(C = 0, 0.5, 1, 1.5, 2, 2.25, 2.5, 2.75, 3, 3.5, 4, 4.5, 5) \times 6$ $(\beta \in [0.825, 0.975]$ with an increment of 0.25). For each combination of (C,β), we compute the value of working for A, $V^A_{i,l,d,t}(C,\beta)$, for each driver i at location l on date d and time interval t, where $d \in ActiveDates_i$ and $t \in [t^s_i + 1$, midnight].

The terminal condition at midnight is that the driver would quit:

$$V^A_{i,l,d,midnight}(\cdot) = 0 \quad \forall l = 1, ..., L$$

We solve backward from midnight, for every 20 minutes, until we reach 20 minutes after the actual time the driver started t^s_i. The hourly financial offer for driver i at time t of day d is given by $w_{i,d,t}$. Define $b(t)$ as the time block that the time interval t is in and $wk(d)$ as an indicator whether the date d is a weekday or weekend. We omit the subscript for time block and day of week for simplicity.

$$\begin{aligned}
V^A_{i,l,d,t}(C,\beta) &= (w_{i,d,t}/3) - C + R^V_l \left(\sum_{k=1}^{L} \pi^V_{l,k} \beta^{\tau^V_{l,k}} \left((\tau^V_{l,k} - 1)(w_{i,d,t}/3 - C) + V_{i,k,d,(t+\tau^V_{l,k})} \right) \right) \\
&\quad + \beta(1 - R^V_l) V_{i,l,d,(t+1)} \\
&= (w_{i,d,t}/3) - C + R^V_l \cdot \\
&\quad \left(\sum_{k=1}^{L} \pi^V_{l,k} \beta^{\tau^V_{l,k}} \left((\tau^V_{l,k} - 1)(w_{i,d,t}/3 - C) + \log \left(\sum_{j \in (w,s,q)} \exp \left(V^j_{i,k,d,(t+\tau^V_{l,k})} \right) \right) \right) \right) \\
&\quad + \beta(1 - R^V_l) \log \left(\sum_{j \in (w,s,q)} \exp \left(V^j_{i,l,d,(t+1)} \right) \right),
\end{aligned}$$

where $V^w_{i,l,d,t}(\cdot) = V^A_{i,l,d,t}(\cdot), V^s_{i,l,d,t}(\cdot) = V^B_{l,b(t),wk(d)}(\cdot)$, and $V^q_{i,l,d,t}(\cdot) = 0$.

2.4.2. Adversarial Estimation for Population Parameters

Next, we present the estimation method for our population parameters. The key challenges to the estimation are that traditional likelihood estimation is infeasible and our model consists of sequential decision-making of a large number of drivers. Simulation-assisted estimation methods are a common alternative. However, they still suffer from the curse of dimensionality in models with rich heterogeneity such as ours and the convergence is not guaranteed or can be extremely slow. Therefore, we adopt a machine-learning-based estimation method called *adversarial estimation* recently proposed by Kaji et al. (2020). We also consider an alternative method, approximate Bayesian computation, that is more in line with other common simulation methods and describe our implementation in Appendix **??**.

Adversarial estimation. Adversarial estimation method is inspired by a machine learning algorithm called a generative adversarial network (GAN). The purpose of GANs is to generate artificial data (e.g., images) that look real. In a similar manner, we can use GANs to obtain estimates of our structural model that can generate data (e.g., decisions of ride-hailing drivers) as if it was produced by real human decision-makers. The idea of GANs and adversarial estimation is the result of a two-player minimax game. One player is a discriminator that evaluates the data and determines whether the data is real or simulated, while the other player is a generator that produces simulated data and is trained to increase the error rate of the discriminator. In our setting, the discriminator is a neural network that acts as a classifier (between real versus simulated data) and the generator is the structural model we aim to estimate. The adversarial estimation framework consists of two levels of estimation. The inner maximization problem is the maximum-likelihood estimator of the discriminator model. Then, the outer minimization problem looks for the parameter values for which real and simulated data are indistinguishable. Kaji et al. (2020) provides theoretical guarantees and characterizes the statistical properties of a GAN-based estimator. In particular, the estimator has the same asymptotic distribution to the optimally weighted simulated method of moments.

Implementation. The outcome data of interest is the fractions of ride-hailing drivers on the focal firm A leaving the platform at different locations at each hour on each day. Let $f_{l,h}^d$ be the fraction of drivers who worked for A on day d and left the platform at location l and hour h. The fraction is for each day, therefore, $\sum_{l \in L} \sum_h f_{l,h}^d = 1$ for any d. For example, each row in the data X would contain the day of week (e.g., Wednesday), the hour (e.g., 9-10am), and 20 columns of the fractions of drivers quitting at one of the 20 locations in New York City.

For a set of parameters $\theta = (\beta, \mu_C, \sigma_C)$, we simulate the outcomes 1,000 times and obtain the simulated data X^θ. We then combine the simulated data X^θ with the actual data X and randomize the order of the rows. For each row, we set the label Y to be 1 if the row is from the real data or 0 if it is from the simulated data. We use a feed-forward neural network as our discriminator D. Our neural network is specified with sigmoid activation functions with three hidden layers: the first with 20 nodes, the second 10 nodes, and the third 1 node, and 10% dropout rate during training. We use the `tensorflow` and `keras` packages in R and the default Adam optimization algorithm. After we train $\hat{D}$ using X and X^θ, we then fit it back to the same inputs to predict the labels $\hat{Y}$. We consider a cross-entropy loss function:

$$Loss = \sum_{i \text{ s.t. } Y_i=1} (-\log(\hat{Y}_i)) + \sum_{i \text{ s.t.} Y_i=0} (-\log(1-\hat{Y}_i))$$

The advantage of using the off-the-shelf neural network is that we can obtain the gradients using automatic differentiation. Given $\hat{D}$, we compute numerical gradient of the loss function with respect to the input data $\{X, X^\theta\}$. Using the chain rule, we obtain the gradient with respect to each of the parameters $\Delta(\theta)$. This step allows us to update $\theta^{(t)}$ to $\theta^{(t+1)}$ with one-step gradient descent: $\theta^{(t+1)} = \theta^{(t)} - \xi\Delta(\theta^{(t)})$, where $\xi > 0$ is a learning rate. We repeat the procedure until we obtain convergence (e.g., $\Delta(\theta) \approx 0$).

2.5. Estimation Results

We present below our estimation results obtained by implementing the adversarial estimation strategy described in Section 2.4.2.

Table 7: Population-level parameter estimates

β	μ_C	σ_C
0.92	1.3371	0.9035

2.5.1. Population-level Parameters

We obtain the discount factor estimate to be 0.92. This estimate is relatively high but reasonable as we consider a 20-minute time interval in our finite horizon. We also find that the population distribution of cost is characterized by the mean $\mu_C = 1.3371$ and the standard variation $\sigma_C = 0.9035$. The average cost of \$1.3371 is in line with the 2014 AAA report that suggests the hourly cost for a sedan is \$4.4 and \$5.48 for SUV, translating into \$1.47 and \$1.83 for a 20-minute interval.

2.5.2. Individual-level Parameters

Now that we have estimated population-level parameters, we determine where in the distribution of costs a particular driver lies by conditioning on her past choices.

Figure 10: Distribution of individual-level cost estimates

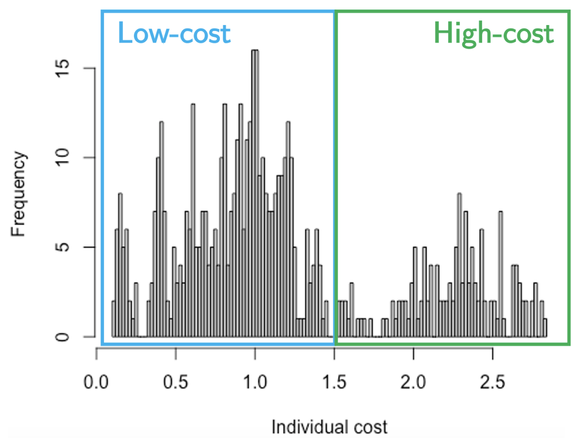
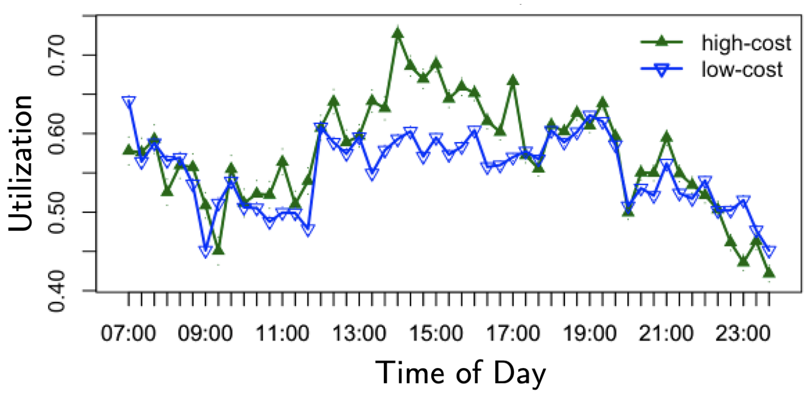


Figure 10 illustrates the distribution of individual-level cost estimates conditional on past decisions. We observe that there are two clusters of drivers' costs: low cost (LC) and high cost (HC). The LC drivers consist of mostly drivers of smaller vehicles. The majority of their starting locations are in Harlem and regions outside of Manhattan. On the other hand, the HC drivers are made up mostly by those driving a larger vehicle. They also tend to start working on the focal platform in Upper or Lower Manhattan. Both groups of drivers work for approximately the same amount per day on average. However, HC drivers experience a significantly higher utilization than LC drivers by approximately 4%. Figure 11 shows the average utilization rate of both types throughout an average weekday. This finding seems to suggest that HC drivers are more strategic in choosing where to start working.

Figure 11: Utilization rate by driver type and time of day



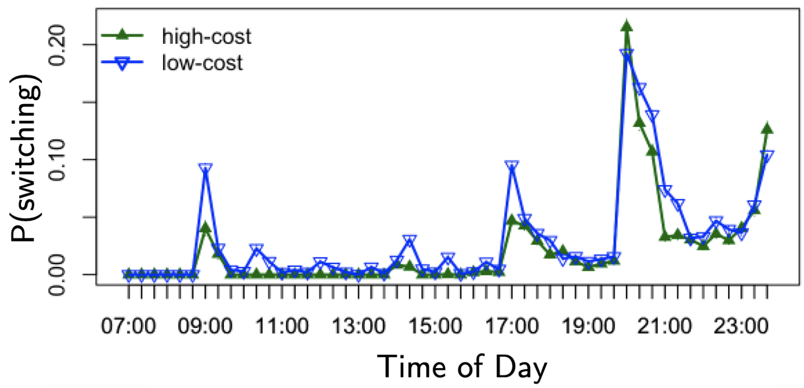
2.5.3. Multihoming Behavior

We find that 40.78% of drivers always work for both platforms every time they work while 14.72% of drivers never switch to work for the competitor platform. We also find that LC drivers are more likely to switch. For both types of drivers, there is an association between facing idleness (e.g., not getting matched to a customer) and switching to work for the competitor platform. Comparing the average probability of switching at each time of day in Figure 12 to the utilization rate illustrated in Figure 11, we observe that the peaks of the switching probability appear to coincide wit the drops of the utilization rate. Interestingly, drivers on the focal platform are more likely to switch when idling even though they are guaranteed an hourly pay. There are two potential reasons. First, switching to the competitor could help them transition to a busier region, potentially increasing future expected earnings. Second, it is possible that service providers such as ride-hailing drivers enjoy non-monetary utility from serving customers.

2.6. Counterfactual Analyses

Structural estimation allows us to quantify the importance of different mechanisms and evaluate counterfactual policies. We consider two levels of policies for our counterfactual analyses, one at the firm level and another at the city level.

Figure 12: Probability of switching to the competitor firm by driver type and time of day

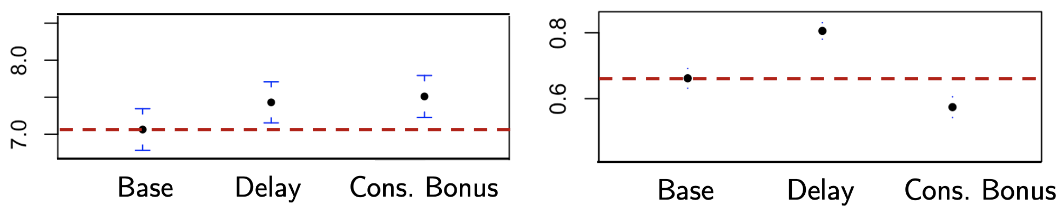


2.6.1. Platform's Policies to Control Multihoming

In 2.5.3, we show that the magnitude of multihoming behaviors among this set of drivers on the focal platform is fairly large. The natural follow-up question is: what can the platform do to control or influence multihoming behavior among drivers? For example, as cities like New York City and Seattle have been capping the number of ride-hailing drivers who can work or licenses to grant, ride-hailing platforms now face with the struggle to encourage drivers from a limited pool to work for them. Multihoming could mean that workers will spend less time on each platform on average, therefore the platforms may prefer to decrease multihoming behavior among workers. On the other hand, for the platforms that compensate workers using a guaranteed pay scheme similar to our industry partner, having too many active workers during the period of low demand could be costly.

We consider two policies that the platforms can implement. First, the *consecutive work bonus* policy requires workers to work consecutively and meet a specified threshold to earn a bonus. This policy is widely used in practice, for example, Uber regularly offers a Consecutive Trips promotion for drivers to make additional earnings by completing multiple trips in a row when and where they expect high demand. Kabra et al. (2017) also finds that threshold incentives are more effective at motivating Singaporean ride-hailing drivers to work than linear incentives. For our counterfactual analysis, we consider a policy in which drivers will only earn their hourly earnings if they stay active for an hour. We also consider the *quit delay* policy in which drivers will face a time delay before they can stop working. On several gig platforms, workers cannot quit right when they want to. They either have to go through multiple confirmation steps before being able to quit or are assigned at least one final assignment. For our counterfactual analysis, we impose a 20-minute delay between the time the driver requests to leave and the time they can actually leave. During the time delay, the driver is still paid a prorated wage for the time and may be assigned to one final trip. If the final trip lasts longer than 20 minutes, the driver will be able to leave right after dropping off the last passenger.

Figure 13: Simulated outcomes across platform-level policies



(a) Work duration on the focal firm in hours (b) Probability of switching to the competitor

Results. We consider four outcomes from this counterfactual analysis: utilization, daily earnings, work duration, and probability of multihoming. The baseline is the current practice without either type of policies. We first find that compared to the baseline the utilization rate when either of the policies is implemented generally drops but not at a statistically significant level. On the other hand, drivers' daily earnings increase compared to the baseline, again not at a statistically significant level.

Figure 13a reports the work duration on the focal firm among drivers under different policies. In terms of work duration, both policies induce the drivers to spend a significantly longer time on the focal platform. Drivers receiving the consecutive work bonus work slightly longer than those facing the time delay. Figure 13b illustrates the probability of drivers on the focal platform switching over to work for the competitor firm within the same day, implying the multihoming behavior. We find that the baseline level of switching is relatively high as close to 70%. The time delay policy significantly increases the switching probability further to over 80%. However, the consecutive work bonus policy leads to a result of a opposite direction. Drivers appear to have a significantly lower probability of switching under the bonus policy.

Our results provide two interesting insights. During peak time when the platform needs to increase its service capacity and encourage workers to stay, using the consecutive work bonus can be effective at lowering the switching probability and increasing work duration on the platform. This could be because such incentive creates a goal or a sense of commitment to the workers. Several gig economy platforms have indeed utilized the threshold policy and incorporated gamification to retain their workers. On the other hand, during the low demand when the platform does not need too many workers or leave their workers idle, imposing a time delay before quitting can be an effective and cost-free policy. By forcing the workers to wait before being able to leave, it triggers them to plan ahead. As we have seen earlier that the drivers appear to be more likely to switch when being idle, we can infer that they do not want to be stuck in the low-demand region. The time delay means that the drivers have to be stuck for a longer period of time. If the drivers want to avoid being stuck in idleness, they have to decide sooner under the time delay policy and thus are more likely to leave.

2.6.2. Policy Analysis: Driver Income Rules

In December 2018, New York City's TLC launched new rules. Drivers should earn at least $17.22 per hour of working and must be paid at least $1.088 x (number of miles traveled) + $0.495 x (trip duration in minutes) for each trip. Ride-hailing platforms such as Uber and Lyft promptly responded to the new rules by restricting driver access to their platforms (e.g., locking a number of drivers out during certain times and/or around certain locations). Juno, a smaller firm, quitted a few months after the new rules took effect. Such effect is in line with the theoretical prediction by Asadpour et al. (2019). The authors show that a minimum wage-type policy is only feasible for loose labor markets. If the minimum wage is too high, the optimal response for the firm is to cap its supply.

The City also provided an estimate that drivers should be paid additional \$172 million in the first four months. However, drivers have reported that the benefit was not substantial. Furthermore, they reported experiencing longer idle time as a result of a larger number of active drivers, driven by the larger guarantee minimum wage. In this counterfactual analysis, we demonstrate that our structural model can help the policymakers better predict the outcome of a new policy.

We consider four scenarios for the policy analysis. First, we consider the *Pre-2019* scenario in which there is no minimum wage policy. Based on our earlier analysis on the data from 2016 to 2017, we have already estimated the model of behavior for this scenario. When the rules took effect in early 2019, they were applied to all ride-hailing platforms except our industry partner. Therefore, we consider the *Current* scenario in which the competitor platform has to follow the rules while the focal platform does not. Comparing this scenario with the Pre-2019 one, we will be able to offer a prediction of the impact of the policy and compare to the actual outcome. We then consider two hypothetical scenarios: *Reverse* in which the rules are only applied to the focal platform but not the competitor, and *Universal* in which the rules are applied to both platforms. We make one assumption that the fares charged to the customers are not affected by the new rules. A report by Parrott and Reich (2018) finds that the fare increase was not significant (e.g., only three to five percent higher).

Results. Our prediction for the Current scenario is that workers would earn 3.5% smaller earnings and spend 2.1% more time idling. This is in line with the actual consequences of the policy that drivers do not generally earn more money and would end up spending more time without work assignment due to an influx of drivers seeking higher earnings through a minimum wage. For the hypothetical scenarios, we first find that the Reverse scenario leads to workers earning 3% higher earnings while spending 3.3% more time idling. The Universal scenario would lead to workers earning 1.2% smaller earnings and spending 7% more time idling. These results suggest that the universal policy might not be optimal as it leads to both lower income and longer idle time. Instead, the policymakers should take into account of how different platforms structure their incentive scheme in order to offer a more effective regulation that benefits all stakeholders.

2.7. Concluding Remarks

With the flexibility in the choice of service, gig workers often work for more than one platform and exhibit a multihoming behavior. An increase in the number of available options has resulted in increased competition among platforms to win over a limited mutual pool of workers. Such competition has only been further exacerbated in cities like New York and Seattle, which recently passed caps on the number of ride-hailing drivers. How workers respond to platform competition is therefore an important topic to study, but studying multihoming behavior empirically is challenging due to the unobservability of workers' options.

In this work, we leverage proprietary data from our ride-hailing industry partner and the publicly available trip record data to develop and estimate a structural model of gig workers' sequential dynamic decisions in the presence of alternative work opportunities. Our major contributions are in the modeling and estimation of dynamic decisions with temporal and spatial components and dynamic outside options, and the development of an efficient simulation-assisted machine learning-based estimation framework in the presence of analytically or computationally intractable likelihood functions and high-dimensional data.

Our results characterize workers' forward-looking behavior and heterogeneous cost of working. We find that workers are strategic in their choice of initial service location to ensure high utilization and are prone to multihoming behavior when facing longer idle times. Our follow-up counterfactual analyses demonstrate the effectiveness of different strategies commonly used in practice and offer insights that can help the firm retain workers during high demand using a consecutive work bonus or nudge them to leave the platform when demand is low by imposing a time delay when the worker requests to leave.

CHAPTER 3 : Learning Best Practices: Can Machine Learning Improve Human Decision-Making?

3.1. Introduction

Workers often spend a significant amount of time on the job learning how to make good decisions that improve their performance (Chui et al. 2012). The impact of a current decision can be highly stochastic and affect future decisions/rewards, making it difficult for them to evaluate the quality of a decision. This issue is further exacerbated by the fact that multiple decisions are often made sequentially, making it hard to determine which decisions are responsible for good outcomes. Many jobs require sequential decision-making; for example, doctors making decisions to optimize the long-term outcomes of their patients (Kleinberg et al. 2015) or drivers on ride-hailing platforms optimizing their long-term profits (Marshall 2020). As a concrete example, physicians seek to learn good strategies for ordering lab tests, since obtaining the appropriate testing results in a timely fashion is necessary to minimize delays in patient visits. Song et al. (2017) finds that experienced physicians have learned to order these tests early on to avoid delays. Despite the simple description of the strategy—"order lab and radiology tests as early in the care delivery process as possible"—learning it on the job is difficult because the connection between when the tests are ordered and the overall quality of care is highly stochastic, and is influenced by other decisions made by the physician as well as unrelated environmental factors such as hospital congestion.

The need to spend time learning on the job has consequences for service quality, since workers likely make suboptimal decisions during this time. For instance, when surgeons first use new devices, surgery duration increases by 32.4% (Ramdas et al. 2017). Thus, whenever possible, workers seek alternative ways to acquire best practices on decision-making. Continuing our example on physician decisions for lab testing, Song et al. (2017) finds that physicians can learn strategies for reducing service time from their better-performing colleagues. This approach is effective precisely because the strategy is simple and easy to communicate, yet time-consuming to discover independently. However, learning from their peers is not always an option for workers; for instance, some workers are comparatively isolated—e.g., physicians working in rural hospitals or operating their own practices or independent workers in the gig economy. In these cases, workers must wastefully spend time independently rediscovering best practices that are already known to their colleagues.

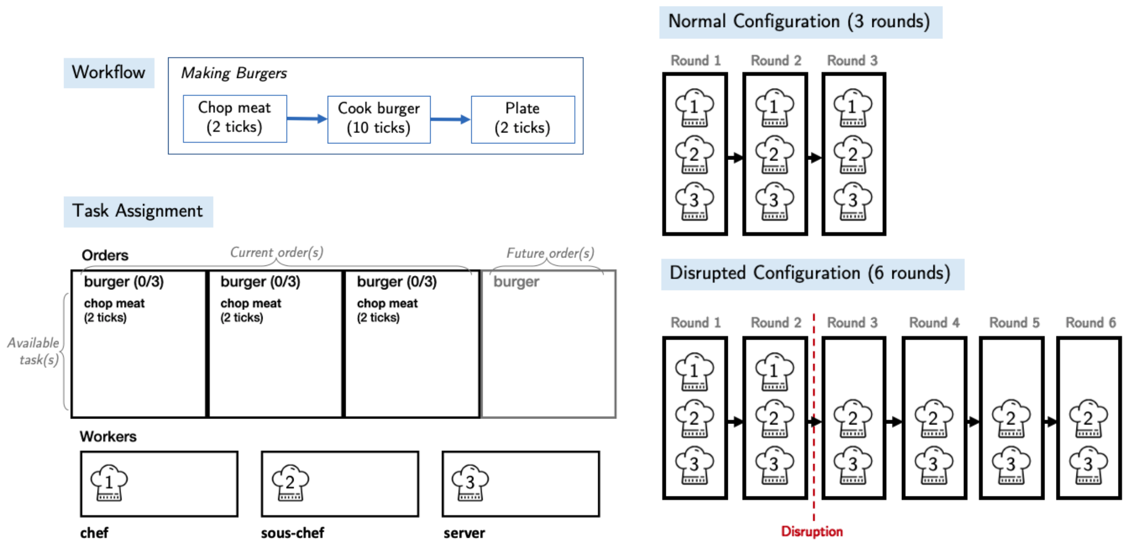
A natural question is whether we can *automatically* discover best practices and convey them to workers to help them improve their performance. In particular, over the past two decades, many domains have accumulated large amounts of *trace data* on human decisions. For example, nearly every physician action is logged in electronic medical record data; every movement of a driver is recorded on a ride-hailing platform; even retail manager decisions on pricing and inventory management are recorded on a daily basis. This data implicitly encodes the collective knowledge acquired by numerous workers about how to effectively perform their jobs. Thus, we might hope to leverage tools from machine learning to mine this data and automatically discover insights that can be used to help workers improve their performance.

In this paper, we study whether machine learning can be used to infer rules that help improve workers' performance at sequential decision-making tasks. In particular, we propose a novel algorithm for mining useful rules or best practices. Our algorithm automatically learns a decision-making rule that, if correctly followed by the human worker, most improves their performance. It does so in two steps. First, our algorithm uses imitation learning (Abbeel and Ng 2004) to learn a model of the current strategy employed by the human workers. These algorithms are designed to reverse-engineer the strategy employed by humans based on data encoding the actions they take in various states. In particular, we use *Q-learning* (Watkins and Dayan 1992) to learn a neural network, called the *Q-network*, that approximates the long-term value of the actions taken by the human workers. In addition to encoding the strategy of the human worker, the Q-network also encodes how changes to the human strategy affect their performance. Then, our algorithm leverages the Q-network to learn a decision-making rule that modifies the human worker strategy in a way that most improves their performance. We must carefully design the search space of decision-making rules so that human workers can correctly follow the rule. That is, the rule must be an *interpretable model* whose computation process can be understood by humans (Letham et al. 2015). In particular, we design the search space to consist of if-then-else rules, and use an approach based on *interpretable reinforcement learning* (Bastani et al. 2018) to learn the best rule. Importantly, despite their simplicity, these if-then-else rules can capture useful insights that are challenging for humans to learn by themselves due to the sequential nature of the decision-making problem.

As a case study, we have designed a game where human participants act as managers for a virtual kitchen. An illustration of this task is shown in Figure 14a. In this environment, the human is shown a set of food orders (e.g., burgers, tacos, etc.), each of which is decomposed into a set of subtasks (e.g., chopping, cooking, serving, etc.). To complete an order, the human must assign each subtask to one of the available virtual workers (e.g., chef, server, etc.). The goal is to do so in a way that completes all the orders as quickly as possible. There are two aspects of this environment that make it challenging: (i) each virtual worker has different skills (e.g., the chef cooks quickly but serves slowly), and (ii) the subtasks have dependencies (e.g., the food must be cooked before it is served). As a consequence, the human must balance leveraging the strengths of each virtual worker (i.e., avoid assigning suboptimal subtasks that the virtual worker is slow to complete) and ensuring that none of the workers are idle (i.e., assign suboptimal subtasks to avoid idling the virtual worker). This environment can be thought of as a networked queuing model with heterogeneous servers—i.e., the subtask dependencies are encoded by the network structure and the virtual workers are the heterogeneous servers.

We conduct a behavioral study using Amazon Mechanical Turk (MTurk) workers to test whether our machine-learning algorithm can learn rules that help human workers improve their performance. Our study is based on two different configurations of our virtual kitchen environment. In the "normal" configuration, the MTurk worker plays three identical instan-

Figure 14: Overview of behavioral study: virtual kitchen management.



(a) Workflow and main experimental task

(b) Study configurations

tiations of the environment. In the "disrupted" configuration, the first two instantiations of the environment are identical to the ones in the normal configuration, but the remaining four instantiations are modified so that a key worker (namely, the chef) is no longer available. These two configurations are visualized in Figure 14b. The disrupted configuration is particularly challenging for the MTurk workers, since they must "un-learn" preconceived notions about the optimal strategy acquired during the first two instatiations. For each of these configurations, we leverage our algorithm to learn interpretable decision-making rules, and then demonstrate how providing this decision-making rule improves the performance of the MTurk workers. Our results show that (i) the tips inferred from our algorithm are effective at significantly improving performance and speeding up learning, (ii) they outperform the tips generated either by previous participants or the baseline algorithm by a significant margin, and (iii) they induce the participants to discover additional optimal strategies beyond what is stated in the tips.

3.1.1. Related Literature and Contributions

Process improvement has always been one of the major emphases both in the operations management literature and in practice. Our work focuses on process improvement from the perspective of individual workers. Scholars have identified various difficulties associated with learning to improve performance. When first experiencing a new work environment, workers tend to have difficulty adjusting, resulting in various degrees of undesirable performance. For instance, as mentioned earlier, Ramdas et al. (2017) finds that when surgeons first use a new surgical device, surgery duration increases by 32.4%, hurting both their service quality and productivity. Bavafa and Jónasson (2020a) shows that unexpected critical medical incidents slow down the ambulance activation among paramedics. The situation exacerbates when inexperienced workers lack a guideline on how to manage their workflow as their prioritization could often be suboptimal and detrimental to productivity (Ibanez et al. 2017). The complex nature of workflow also plays a role. Workers tend to focus on immediate challenges and ignore opportunities for learning (Tucker et al. 2002) and switching between tasks could hurt as much as 20% of their productivity (Gurvich et al. 2019). In many collaborative work settings, productivity depends on one's co-workers. Collaboration is particularly challenging in distributed work, where there is considerable uncertainty about others' behaviors (Weisband 2002, Mao et al. 2016). This is especially true in healthcare, where delivery processes involve numerous interfaces and patient handoffs among multiple healthcare practitioners with varying levels of training and prior experiences working together (Hughes et al. 2008, Akşin et al. 2020).

To increase reliability and reduce process variation, process standardization is commonly implemented to form best practices (Nonaka and Takeuchi 1995, Pfeffer et al. 2000, Spear 2005). Process standardization is generally a two-step process: creating the standards and then communicating them. Creating standards and developing knowledge of best practices are known to be hard as they take time (Nonaka and Takeuchi 1995) and knowledge transfer often fails across organizational borders (Szulanski 1996, Argote 2012). A rich literature in operations management and organizational behavior has shown how various aspects of experiences can improve individuals' productivity and performance. For example, professional web developers frequently learn new concepts and strategies by trial and error (Dorn and Guzdial 2010). Past experiences on the same or related tasks, even subtasks, have a significant effect on performance (Huckman and Pisano 2006, Kc and Staats 2012), a variety of experiences could hinder workers' ability to identify best practices (Kc and Staats 2012). Furthermore, Bavafa and Jónasson (2020b) shows that greater prior experience reduces variance of performance. Social interaction is another common way to learn. Therapy workers learn from clients' feedback to adjust their treatment process (Brattland et al. 2018). Workers also learn significantly from their colleagues, particularly those with a high level of knowledge or valuable skills (Herkenhoff et al. 2018, Jarosch et al. 2019). Song et al. (2017) shows that by publicly disclosing relative performance feedback, physicians can better identify their top-performing co-workers, enabling the identification and validation of best practices. Working alongside experienced peers is shown to improve workers' performance (Chan et al. 2014, Tan and Netessine 2019). Team experience and familiarity with one another and with the tasks are associated with both team and individual performance (Akşin et al. 2020, Kim et al. 2020). However, these learning strategies can be inefficient as they rely on the availability of experts and knowledge of best practices. Given well-documented difficulties in learning on the job and identifying best practices, our work proposes an effective approach to automatically extract best practices from logs of historical decisions and outcomes.

Besides identifying best practices, effectively sharing and encouraging workers to adopt them are known to be challenging (Tucker et al. 2007). One way to improve such knowledge transfer is to structure it as a simple rule. The clarity of simple rules allows workers to gain deeper understanding of the environment and potential improvement (Sull and Eisenhardt 2015, Gleicher 2016). A simple training intervention is also shown to improve decision-making by persistently reducing cognitive biases (Morewedge et al. 2015, Sellier et al. 2019). Thanks to the fast-growing advancement of artificial intelligence, machine-learning models have demonstrated great success in learning complex systems and making predictions that help guide high-stakes decision-making in various domains, from healthcare to criminal justice. However, most commonly used black-box models do not provide users with transparency, accountability, or explanations. The lack of human understanding of how algorithms work poses serious problems to society (Rudin 2019) and leads to aversion to adopting these tools (Dawes et al. 1989, Dietvorst et al. 2015). In recent years, significant efforts have been dedicated towards the development of models that are inherently interpretable (e.g., see Murdoch et al. (2019) for an in-depth review of methods and applications of interpretable machine learning). Incorporating human domain knowledge into algorithms has also received increased attention recently (Arvan et al. 2019, Ibrahim et al. 2020). Our work contributes to this stream of literature in two ways. First, we show that a simple intervention–providing a simple tip–can help improve worker performance over time and speed up their learning process. Second, we develop a novel algorithm that leverages the largely untapped potential of worker trace data to complement existing training programs and learning among workers.

3.2. Learning Interpretable Tips for Improving Human Performance

Consider a human making a sequence of decisions to achieve some desired outcome. We study settings where current decisions affect future outcomes—for instance, if the human decides to consume some resources at the current time step, they can no longer use these resources in the future. These settings are particularly challenging for decision-making due to the need to reason about how current actions affect future decisions, making them ideal targets for leveraging tips to improve human performance. In particular, our goal is to provide insights to the human that enable them to improve their performance. In this section, we describe our algorithm for computing rules designed to improve the performance of human workers. We begin by formalizing the tip inference problem, and then describe our algorithm for solving this problem.

3.2.1. Background on MDPs

Because we are focused on sequential decision-making, we assume the decision-making problem is modeled by a Markov Decision Process (MDP). In particular, consider an

MDP $\mathcal{M} = (S, A, P, R, \gamma)$, where S is a finite set of states, A is a finite set of actions, $P : S \times A \times S \to \mathbb{R}$ encodes the transition probabilities (i.e., $P(s' \mid s, a)$ is the probability of transitioning from state s to state s' upon taking action a), $R : S \to \mathbb{R}$ denotes rewards (i.e., $R(s)$ is the reward obtained in state s), and $\gamma \in (0, 1)$ is a discount factor. In addition, we assume there is a deterministic initial state $s_0 \in S$. Finally, we also assume there is a finite time horizon $T \in \mathbb{N}$ after which the rewards are zero.

Now, given a stochastic policy $\pi : S \times A \to \mathbb{R}$ (i.e., $\pi(a \mid s)$ is the probability of taking action a in state s), a *rollout* in M uses π to generate a random sequence of state-action-reward tuples $\zeta = ((s_1, a_1, r_1), ..., (s_T, a_T, r_T))$, where $a_t \sim \pi(\,\cdot \mid s_t)$, $r_t = R(s_t)$, and $s_{t+1} \sim P(\,\cdot \mid s_t, a_t)$. We use $D^{(\pi)}$ to denote the distribution over rollouts using π. Note that if the MDP transitions P and the policy π are both deterministic, then ζ is also deterministic.

The *cumulative expected reward* of π is

$$J(\pi) = \mathbb{E}_{\zeta \sim D^{(\pi)}} \left[\sum_{t=1}^{T} \gamma^t r_t \right].$$

Finally, the value function $V^{(\pi)} : S \to \mathbb{R}$ and Q function $Q^{(\pi)} : S \times A \to \mathbb{R}$ of π are the unique solutions to the recursive system of equations

$$\begin{aligned} V^{(\pi)}(s) &= \mathbb{E}_{\pi(a|s)}[Q^{(\pi)}(s, a)] \\ Q^{(\pi)}(s, a) &= R(s) + \gamma \cdot \mathbb{E}_{P(s'|s,a)}[V^{(\pi)}(s')], \end{aligned}$$

respectively. Intuitively, $V^{(\pi)}(s)$ is the cumulative expected reward of using π if the initial state is s, and $Q^{(\pi)}(s, a)$ is the cumulative expected reward of using π from state s, but where the first action taken is fixed to be a.

3.2.2. Problem Formulation

Given an MDP and a human acting in that MDP, our goal is to learn a decision-making rule that most improves the performance of the human. In particular, we model the human as executing a *human policy* π_H; we measure their performance as the cumulative expected reward $J(\pi_H)$ that they achieve. To ensure that the decision-making rule can be understood by the human worker, we restrict to learning rules of the form

$$\text{if [state constraint] then [action].}$$

Note that this rule specifies the action to take in a portion of the state space; in the remainder of the state space, the human should continue to make decisions using their own policy π_H. More precisely, assuming we have a mapping $\phi : S \to \{0,1\}^d$ of states to a set of binary properties $\phi(s)_i$, then a state constraint is a predicate $\psi : S \to \{0,1\}$ of the form

$$\psi(s) = (\phi(s)_{i_1} = b_1) \wedge ... \wedge (\phi(s)_{i_k} = b_k).$$

Then, the rule is a pair $\rho = (\psi, a)$ of a predicate ψ and an action a. Intuitively, this rule says to take action a in state s if s satisfies ψ (i.e., $\psi(s) = 1$); otherwise, the human should take an action using their own policy π_H. More precisely, given a stochastic policy π and a rule $\rho = (\psi, a)$, we define the *rule-following policy* $\pi \oplus \rho : S \times A \to \mathbb{R}$

$$(\pi \oplus \rho)(s, a) = \begin{cases} \mathbb{1}(a = a') & \text{if } \psi(s) = 1 \\ \pi(s, a) & \text{otherwise.} \end{cases}$$

In particular, $\pi_H \oplus \rho$ represents the setting where the human worker exactly follows rule ρ—i.e., they use the action recommended by ρ when applicable and use their own policy π_H otherwise.[1] Finally, given a class of rules $\rho \in \mathcal{R}$, our goal is to choose the one that most improves the performance of the human worker—i.e.,

$$\rho^* = \arg\max_{\rho \in \mathcal{R}} J(\pi_H \oplus \rho). \tag{3.1}$$

[1] Although we rank rules assuming humans follow our tips exactly, this is not the case in practice. Nevertheless, our behavioral experiments demonstrate that our tips significantly improve performance relative to other types of tips.

To guide our algorithm for learning ρ^*, we assume we are given an *expert policy* π_* that achieves high performance $J(\pi_*)$. In principle, we can compute the exact optimizer $\pi_* = \arg\max_\pi J(\pi)$ using dynamic programming. However, this approach is computationally intractable for large state spaces. Instead, we can use techniques such as model-free reinforcement learning (Watkins and Dayan 1992, Sutton et al. 2000) to compute π_* that approximately optimizes $J(\pi)$. These approaches rely on our assumption that the MDP structure is known; when it is unknown, our algorithm can instead leverage sampled rollouts from a human expert.

3.2.3. Tip Inference Algorithm

Now, we describe our algorithm for maximizing the objective in (3.1). Ideally, our algorithm would simply enumerate $\rho \in \mathcal{R}$, compute $J(\pi_H \oplus \rho)$, and return the rule ρ that achieves the highest score. The key challenge is how to compute the value of the objective $J(\pi_H \oplus \rho)$ in (3.1) for a candidate tip $\rho \in \mathcal{R}$. First, we leverage the following result from Bastani et al. (2018):

Lemma 3.2.1 *For any policy π, we have*

$$J(\pi) = \mathbb{E}_{\zeta \sim D^{(\pi)}}\left[\sum_{t=1}^{T} Q^{(\pi_*)}(s_t, a_t)\right],$$

We can use this result to rewrite the objective $J(\pi_H \oplus \rho)$ in (3.1). However, we do not have access to samples $\zeta \sim D^{(\pi_H \oplus \rho)}$. To address this issue, we use an approximation where we assume that the distribution over rollouts of the human user is not significantly affected by the rule—i.e., $D^{(\pi_H \oplus \rho)} \approx D^{(\pi_H)}$. Then, we have

$$\begin{aligned} J(\pi_H \oplus \rho) &= \mathbb{E}_{\zeta \sim D^{(\pi_H \oplus \rho)}}\left[\sum_{t=1}^{T} Q^{(\pi_*)}(s_t, a_t)\right] \\ &\approx \mathbb{E}_{\zeta \sim D^{(\pi_H)}}\left[\sum_{t=1}^{T} Q^{(\pi_*)}(s_t, (a_t \oplus \rho)(s_t)\right], \end{aligned}$$

where given rule $\rho = (\psi, a)$ and action a', we define

$$(a' \oplus \rho)(s) = \begin{cases} a & \text{if } \psi(s) = 1 \\ a' & \text{otherwise.} \end{cases}$$

Next, we can approximate this objective using sampled rollouts based on the human policy—i.e., given samples $\zeta^1, ..., \zeta^k \sim D^{(\pi_H)}$, we have

$$\hat{\rho} = \arg\max_{\rho \in \mathcal{R}} \frac{1}{k} \sum_{i=1}^{k} \sum_{t=1}^{T} Q^{(\pi_*)}(s_t^i, (a_t^i \oplus \rho)(s_t^i)).$$

The remaining challenge is that we do not have access to the Q-function $Q^{(\pi_*)}$ of the expert policy π_*. We can learn an estimate $\hat{Q}$ of $Q^{(\pi_*)}$ using supervised learning based on sampled rollouts $\zeta \sim D^{(\pi_*)}$. In particular, given samples $\zeta^1, ..., \zeta^h \sim D^{(\pi_*)}$, we solve the optimization problem

$$\hat{Q} = \arg\min_{Q \in \mathcal{Q}} \sum_{i=1}^{h} \sum_{t=1}^{T} \left(Q(s_t^i, a_t^i) - Q_t^i\right)^2 \qquad \text{where} \qquad Q_t^i = \sum_{\tau=t+1}^{T} r_\tau^i.$$

Here, Q_t^i is an unbiased estimate of $Q^{(\pi^*)}(s_t^i, a_t^i)$. For instance, we could choose $\mathcal{Q}$ to be a random forest or a neural network. Then, our objective becomes

$$\hat{\rho} = \arg\max_{\rho \in \mathcal{R}} \frac{1}{k} \sum_{i=1}^{k} \sum_{t=1}^{T} \hat{Q}(s_t^i, (a_t^i \oplus \rho)(s_t^i)).$$

3.3. Case Study: Virtual Kitchen Management Game

We seek to evaluate whether our algorithm can reliably improve worker performance in a controlled environment. To this end, we have developed a sequential decision-making task in the form of a virtual kitchen game that can be played by individual human users.

In this game, the human user takes a role of a manager of a virtual kitchen. The overall goal is to complete a fixed set of n food orders (e.g., burgers, tacos, etc.), where order

$j \in \{1, ..., n\}$ consists of k_j subtasks (e.g., chopping, cooking, serving, etc.). To complete an order, the human user must assign each of these subtasks to one of the available virtual workers (e.g., chef or server), ideally accounting for the heterogeneous skillset of each worker. The game operates in discrete time steps called *ticks*. On each tick, the human user can assign any of the available subtasks to any of the available virtual workers, where (i) a subtask is available if all its prerequisites have been completed but it has not yet been assigned, and (ii) a virtual worker is available if they are not currently working on another subtask. Importantly, the human user can choose not to assign any subtask to a virtual worker even if they are available—e.g., to strategically wait and assign the worker a more appropriate subtask that will become available at a later tick. On each tick, the human user makes their desired assignments, and then clicks a "next tick" button on the screen; upon clicking this button, the assignments are made and the game is incremented to the next tick. This process repeats until all orders are complete. The goal of the human user is to assign all subtasks to the virtual workers in a way that minimizes the total number of ticks it takes to complete all the orders.

There are two aspects of the game which make it challenging to play optimally. First, the subtasks have dependencies—e.g., the burger must be cooked before it can be served. Second, the virtual workers have different skills that affect how long they take to complete a subtask—e.g., a chef can cook quickly but is slow at serving, whereas a server can serve quickly but is slow at cooking. Thus, the human user must make trade-offs such as deciding whether to greedily assign a worker to a task that they are slow to complete, or leave them idle in anticipation of an upcoming task they can complete quickly.

3.3.1. Formulation as an MDP

At a high level, the states encode the progress towards completing all the food orders, the actions encode the available assignments of subtasks to workers, and the rewards encode the number of ticks the user takes to complete all the orders. More specifically, the states encode the following information: (i) in all the orders, which subtasks have been completed so far, and (ii) which subtask has been assigned to each virtual worker (if any), and how many ticks remain for the virtual worker to complete that subtask. Next, the actions consist of all possible assignments of available subtasks to available virtual workers. Finally, the reward is -1 at each tick, until all orders are completed—thus, the total number of ticks taken to complete all orders is the negative reward.

3.3.2. Search Space of Tips

Next, we describe the search space of tips (i.e., rules) that are considered by our algorithm. Each tip is actually composed of a set of rules inferred by our algorithm. Recall that our algorithm considers rules in the form of an if-then-else statement that says to take a certain action in a certain state. One challenge is the combinatorial nature of our action space—there can be as many as $\frac{k!}{(k-m)!}$ actions, where m is the number of workers and $k = \sum_{j=1}^{n} k_j$ is the total number of subtasks. The large number of actions can make the rules very specific—e.g., simultaneously assigning three distinct subtasks to three of the virtual workers. Instead, we decompose the action space and consider assigning a single subtask to a single virtual worker. To be precise, we include three features in the predicate ϕ: (i) the subtask being considered, (ii) the order to which the subtask belongs, and (iii) the virtual worker in consideration. Then, our algorithm considers rules of the form

$$\text{if (order} = o \ \wedge \ \text{subtask} = s \ \wedge \ \text{virtual worker} = w) \text{ then (assign } (o, s) \text{ to } w),$$

where o is an order, s is a subtask, and w is a virtual worker.

Even with this action decomposition, we found that these rules are still too challenging for human users to internalize since the rules are very specific. Instead, we post-process the rules inferred by our algorithm by aggregating over tuples (o, s, w) that have the same s and w—e.g., instead of considering two separate rules[2]

[2] We experimented with *combinations* of rules in exploratory pilots, and found that MTurk workers were unable to operationalize and comply with such complex tips even though they might be part of an optimal strategy.

$$\text{if (order} = \text{burger}_1 \ \wedge \ \text{subtask} = \text{cooking} \ \wedge \ \text{virtual worker} = \text{chef) then (assign (burger}_1, \text{cooking) to chef),}$$
$$\text{if (order} = \text{burger}_2 \ \wedge \ \text{subtask} = \text{cooking} \ \wedge \ \text{virtual worker} = \text{chef) then (assign (burger}_2, \text{cooking) to chef),}$$

we merge them into a tip

$$\text{assign cooking to chef 2 times.}$$

In other words, a tip is a combination of rules $\rho = (\rho_1, ..., \rho_k)$. The score that we assign to such a tip is $J(\rho) = \sum_{i=1}^{k} J(\rho_i)$. Then, we choose the tip that achieves the highest score.

3.3.3. Rule Inference Algorithm Implementation

First, we describe how we compute the expert Q-function Q^*. In principle, we could use dynamic programming to solve for the optimal value function V^*, and then compute the optimal Q-function based on V^*. However, while our state space is finite, it is still too large for dynamic programming to be tractable. Instead, we use the policy gradient algorithm (which is widely used for model-free reinforcement learning) as a heuristic to learn an expert policy π_* for our MDP (Sutton et al. 2000).

At a high level, the policy gradient algorithm searches over a family of policies π_θ parametrized by $\theta \in \Theta \subseteq \mathbb{R}^{d_\Theta}$; typically, π_θ is a deep neural network, and θ is the corresponding vector of neural network parameters. This approach requires featurizing the states in the MDP—i.e., constructing a feature mapping $\phi : S \to \{0,1\}^d$. Then, the neural network policy π_θ takes as input the featurized state $\phi(s)$, and outputs an action $\pi_*(\phi(s)) \in A$ to take in state s.[3] Then, the policy gradient algorithm performs stochastic gradient descent on the objective $J(\pi_\theta)$, and outputs the best policy $\pi_* = \pi_{\theta^*}$. In general, $J(\pi_\theta)$ is nonconvex so this algorithm is susceptible to local minima, but it is well known that it performs exceptionally well in practice.

[3] To be precise, $\pi_*(\phi(s))$ outputs a probability $\pi_*(a \mid \phi(s))$ for each action $a \in A$ of taking a in state s. We take the action a with the highest probability.

In our implementation, the state features include the availability of each sub-task (for each order), the current status of each worker, and the time index. We take π_θ to be a neural network with 50 hidden units; to optimize $J(\pi_\theta)$, we take 10,000 stochastic gradient steps with a learning rate of 0.001. In addition, since our MDP has finite horizon, we use a discount factor of $\gamma = 1$.

Once we have computed π_*, we use the supervised learning algorithm described in Section 3.2 to learn an estimate $\hat{Q}$ of the optimal policy's $Q^{(\pi_*)}$; specifically, we choose $\hat{Q}$ to be a random forest (Breiman 2001). The random forest operates over the same featurized states as the neural network policy—i.e., it has the form $\hat{Q}(\phi(s), a) \approx Q^{(\pi_*)}(s, a)$.

Finally, we apply our algorithm to inferring rules on state-action pairs collected from observing human users playing our game. Because our goal is to help human users improve, we restrict our data to the bottom 25% of human users in terms of performance. In addition, we apply two additional postprocessing steps to the set of candidate rules. First, we eliminate rules that apply in less than 10% of the states occurring in the human trace data—i.e., the predicate $\psi(s) = 1$. This step eliminates high-variance rules that have large benefits, but are useful only a small fraction of the time. Second, we eliminate rules that, when they apply, disagree with the expert policy more than 50% of the time—i.e., for a rule (ψ, a), $\psi(s) = 1$ and $a \neq \pi^*(s)$. This step eliminates rules that have large benefits on average, but sometimes offer incorrect advice that can confuse the human user (or cause them to distrust our tips).

3.3.4. Baseline Algorithm

A simpler approach is to directly imitate the optimal policy rather than indirectly using the Q function. In particular, given rollouts $\hat{D}$ from a policy π, we compute the frequency of state-action pairs in $\hat{D}$—i.e.,

$$\hat{C}(\psi, a) = \log\left(1 + \#\{(\psi, a) \in \hat{D}\}\right),$$

where $\#\{(s, a) \in \hat{D}\}$ is the number of times the action a was taken when the state constraint ψ was active in one of time steps in a rollout in D^{π}. Then, the baseline algorithm selects the rule "if $\psi(s)$ then a" with the highest $\hat{C}(\psi, a)$. Then, we post-process these tips in the same way as we process the tips inferred using our algorithm. This comparison evaluates the importance of accounting for the reward structure when selecting good tips.

3.4. Experimental Design

We perform a behavioral study to evaluate whether the tips inferred using our algorithm can help human participants improve their performance at complex sequential decision-making tasks—specifically, in our virtual kitchen game. Our goal is to address the following research questions:

- Can our algorithm infer tips that help participants improve their performance on a complex sequential decision-making task?
- How do the tips inferred by our algorithm compare to other approaches, including tips suggested by experienced human participants as well as a baseline algorithm-inferred tip?
- Does the inferred tip help improve human performance solely because the participant follows the tip, or does it induce them to improve in additional ways?

Recall that our tip inference algorithm requires data on the human performance so it can focus on conveying information not already known by the humans. For example, if it is fairly common for human participants to learn not to assign a lengthy cooking task to a virtual server, our algorithm will search for other rules that are less obvious but vital to performance improvement. Thus, our behavioral experiment proceeds in two phases. First, we collect data on the actions taken by the humans participants when they are not provided with any tips, and use our algorithm in conjunction with this data to infer tips. Second, we evaluate whether providing these tips can significantly improve the performance of subsequent human participants, compared both to a control group of participants who are not provided with any tips as well as two other groups of participants who are each provided with tips of different sources (i.e., previous participants or a baseline algorithm).

Each human user plays the game several times in sequence, allowing them to learn good strategies over time. For each time the user plays the game, we need to specify the orders that must be completed as well as the available virtual workers; we refer to this specification as a *scenario*. Furthermore, we refer to the overall sequence of scenarios played by the user as a *configuration*.

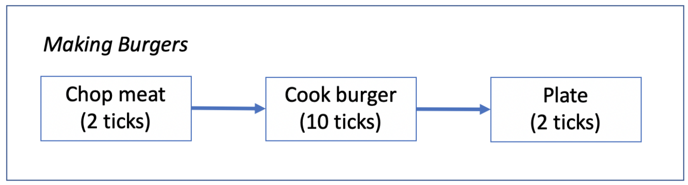
We evaluate our algorithm based on two different configurations of our queueing game that are designed to evaluate different conditions under which tips might be useful. First, the *normal configuration* consists of a single scenario we refer to as the *fully-staffed scenario*. Thus, our goal is to infer tips that help the human users fine-tune their performance at this scenario. Second, the *disrupted configuration* starts with the fully-staffed scenario, but then switches to a modified scenario called the *understaffed scenario*. Intuitively, we expect the human workers to acclimate to the fully-staffed scenario; thus, they may have difficulty adapting to the understaffed scenario where the high-level strategy is very different. Thus, our goal is to infer tips that convey shifts in strategy that are needed to perform well in the new scenario.

For a given configuration, we additionally vary the tip that is shown to the user. Potential tips include: no tip (i.e., the control), our algorithm-inferred tip, a baseline tip, and a human-suggested tip. Our goal is to understand how the choice of tip affects the performance of the human users in the context of that configuration.

3.4.1. Virtual Kitchen Scenarios & Configurations

Our experimental design is based on two scenarios of the virtual kitchen, differing in terms of which workers are available. In the *fully-staffed scenario*, the human user has access to three virtual workers, whereas in the *understaffed scenario*, the human user only has access to two workers. The scenarios are identical in terms of the orders that must be completed. The orders are all the same dish—specifically, they must complete four burger orders. To complete a single burger order requires three subtasks: (i) chopping meat, (ii) cooking burger, and (iii) plating. Each subtask can only be started once the previous one has been completed. The subtasks in the burger order are illustrated in Figure 15.

Figure 15: Subtasks required to make a burger with a median processing time for each subtask.

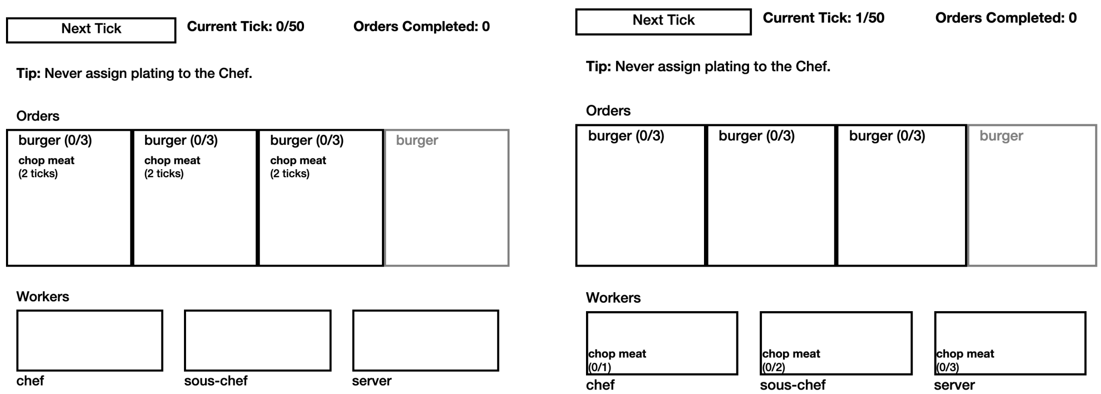


There are three possible virtual workers in the kitchen: chef, sous-chef, and server. The chef is fastest at chopping and cooking, but is slowest at plating. The sous-chef is a "jack-of-all-trades", who can perform all tasks at an intermediate speed. Finally, the server is fastest at plating, but slowest at chopping and cooking. Participants are not given the exact number of ticks each worker takes to complete each subtask. Instead, when a subtask becomes available, participants are shown the median number of ticks required to complete that subtask. The true number of ticks for a given worker is only revealed if they assign the subtask to that worker. The human user must experiment to learn this information by playing the game. Figure 16 shows two screenshots of the game. While the optimal policy in these types of games are typically quite complex and must be solved approximately, we find that a significant number of human users are able to learn very efficient policies by playing the game 3-6 times in a row. For instance, they identify bottleneck subtasks, and learn to mitigate them by assigning these subtasks to their most capable virtual worker.

As discussed above, we consider two different scenarios for our virtual kitchen. First, in the *fully-staffed scenario*, the human user has access to all three virtual workers—i.e., the chef, sous-chef, and server. In contrast, in the understaffed scenario, the user has access to only two virtual workers—namely, the sous-chef and server.

Finally, we describe the two configurations of our game used in our study. First, in the

Figure 16: Example screenshots from the game.



(a) The initial state where participants observe available subtasks from current orders, median times to completion, and three idle virtual workers. The interface also shows the current tick, time limit, current progress, and treatment-specific tip.

(b) The next state after all three previously available subtasks were assigned to the virtual workers and the true completion times were realized, revealing different levels of virtual workers' skills.

normal configuration, the human user plays the fully-staffed scenario for three rounds. Second, in the disrupted configuration, they first play two rounds of the fully-staffed scenario, followed by four rounds of the understaffed scenario.

3.4.2. Experimental Procedure

We perform separate experiments for each configuration of our game. The high-level structure of our experimental design for each configuration is the same; they differ in terms of when we show tips to the user and which tips we show. For each configuration, our experiment proceeds in two phases. Before starting our game, each human user is shown a set of game instructions and comprehension checks; then, they play a practice scenario twice (with an option to skip the second one).[4]

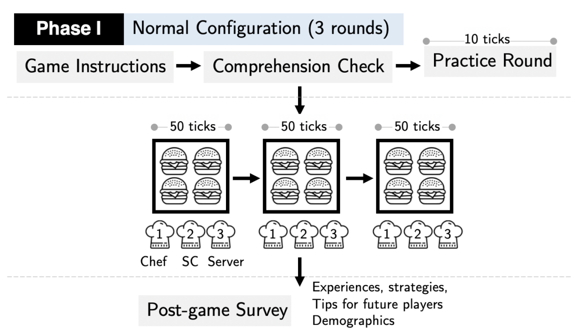
In the first phase, we recruit 200 participants via Amazon Mechanical Turk to play the game. At the end of all rounds of play, we give users a post-game survey where we ask several questions regarding their experience with the game. Additionally, we ask the participants to suggest a tip for future players. In particular, we show each participant a comprehensive

[4]The practice scenario is meant to familiarize participants with the game mechanics and the user interface. In this scenario, they manage three identical chefs to make a single, simple food order. This food order is significantly different than the burger order used in the main game.

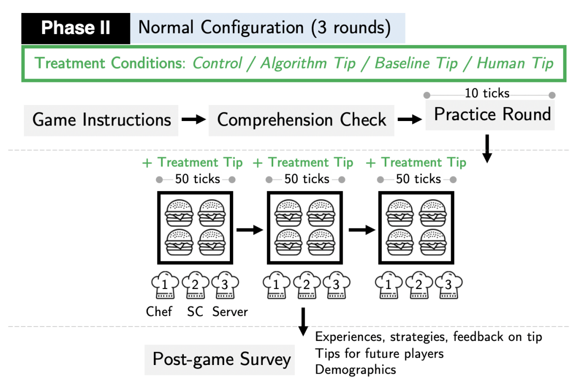
list of candidate tips and ask them to select the one they believe would most improve the performance of future players. This list of tips is constructed by merging three types of tips: (i) all possible tips of the format described in Section 3.3.2 (e.g., “Chef shouldn’t plate.”), (ii) generic tips that arise frequently in our exploratory user studies (e.g., “Keep everyone busy at all time.”), (iii) a small number of manually constructed tips obtained by studying the optimal policy (e.g., “Chef should chop as long as there is no cooking task”). Importantly, this list always contains the top tip inferred using our algorithm.

Next, we use the data from the final round played by the participants to infer tips in three ways. First, we train our tip inference algorithm described in §3.3.3 on the experimental data and obtain an *algorithm tip*. Second, we implement the baseline algorithm described in §3.3.4, which directly tries to imitate the optimal policy based on the frequency of state-action pairs observed in the experimental data, and yield an *baseline tip*. Finally, we take the tip with the most votes from the participants in the post-game survey as a *human tip*.

Figure 17: Study flow for the normal configuration.



(a) Phase I



(b) Phase II

In the second phase, we evaluate the effectiveness of the each of the inferred tips based on the data from the first phase. In this phase, human users are randomly assigned to one of four arms, which differ in terms of the tip that is shown to the user. These arms include the *control arm* (i.e., no tip), the *algorithm arm* (i.e., the tip inferred by our algorithm), the *baseline arm* (i.e., the tip inferred by the baseline algorithm), and the *human arm* (i.e., the tip most frequently recommended by human users). We recruited 350 MTurk users to play each arm in each configuration, totaling 2,800 users.

The specific tips we show on each round depends not just on the arm, but also varies from round to round depending on the configuration. For the normal configuration (i.e., the human user plays the fully-staffed scenario for three rounds), we show the tip for the current arm on all three rounds. Figures 17a and 17b illustrate the study flow for the normal configuration. However, for the disrupted configuration configuration (i.e., they first play two rounds of the fully-staffed scenario, followed by four rounds of the understaffed scenario), the tip for the current arm is specific to the understaffed scenario. Thus, we only show the tip for the current arm on rounds 3-6. In all arms, for rounds 1 and 2, we show the tip inferred using our algorithm for the fully-staffed scenario from the normal configuration. By doing so, we help the human users learn more quickly to play the fully-staffed scenario before switching to the understaffed scenario. Figures 18 and 19 illustrate the study flow for the disrupted configuration.

Figure 18: Study flow for Phase I of the disrupted configuration.

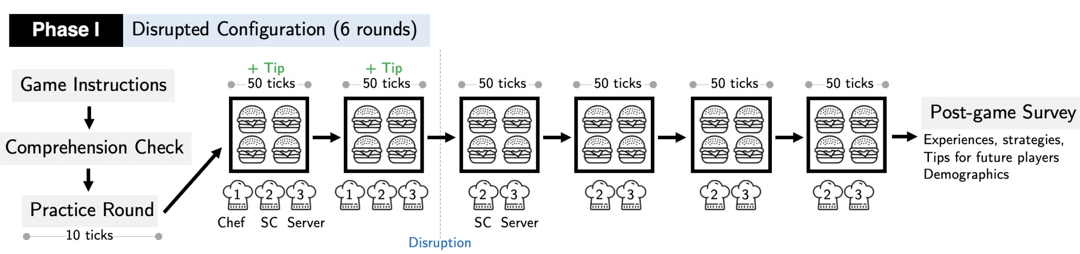
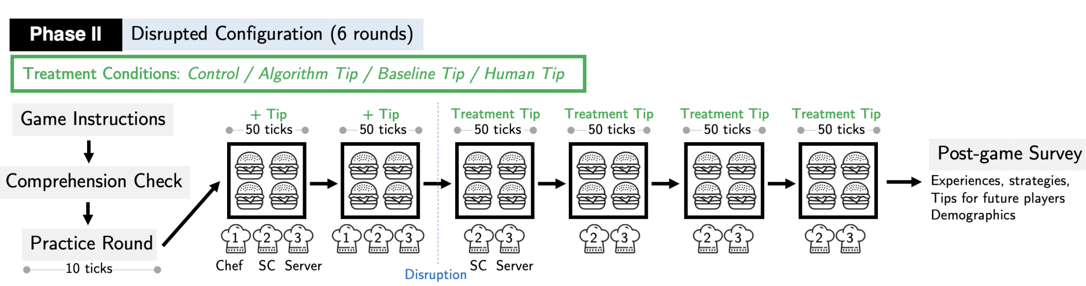


Figure 19: Study flow for Phase II of the disrupted configuration.



Each participant receives a participation fee of $0.10 for each round they completed. We also provide a bonus based on their performance, measured by the number of ticks taken to complete each round. The bonus ranges from \$0.15 to \$0.75 per round. The full pre-registration document for our study is available at `https://aspredicted.org/blind.php?x=8ye5cb`.

3.4.3. Hypotheses

We are interested in addressing three sets of hypotheses with our experiment.

1. Do humans perform better with the tip from our algorithm compared to receiving no tips?

   H1a: In the normal configuration, participants who receive the tip generated by our algorithm will perform better than those not receiving any tips.

   H1b: In the disrupted configuration, participants who receive the tip generated by our algorithm will perform better than those not receiving any tips.

2. Do humans perform better with a tip from our algorithm compared to the tip most frequently suggested by other humans who have completed the game?

   H2a: In the normal configuration, participants who receive the tip generated by our algorithm will perform better than those receiving the tip most frequently suggested by past participants who also played the normal configuration.

   H2b: In the disrupted configuration, participants who receive the tip generated by our algorithm will perform better than those receiving the tip most frequently suggested by past participants who also played the disrupted configuration.

3. Do humans perform better with the tip from our algorithm compared to the tip from a baseline tip mining algorithm?

   H3a: In the normal configuration, participants who receive the tip generated by our algorithm will perform better than those receiving the tip generated by the baseline.

H3b: In the disrupted configuration, participants who receive the tip generated by our algorithm will perform better than those receiving the tip generated by the baseline.

To do so, we perform six two-sample one-sided t-tests to compare the distributions of number of ticks to completion of the final round of each configuration. We also consider additional secondary outcome measures including the learning rate (e.g., performance across rounds), the fraction of participants who completed each round of the game by taking the optimal number of steps, and how well the participants complied with the provided tip and learned additional optimal strategies.

3.5. Experimental Results

We show that our tips significantly improve participant performance compared to the control group, as well as compared to two baselines: (i) a baseline algorithm that naïvely tries to match the optimal policy, and (ii) tips suggested by previous participants. These results demonstrate that despite their simplicity and conciseness, our tips capture strategies that are hard for participants to learn and can significantly improve their performance. In addition, we find evidence that the participants were not blindly following our tips, but combine them with their own experience to improve performance. Finally, we also find evidence that participants build on our tips by discovering additional strategies beyond the ones stated in the tips. We first discuss the tips inferred from our algorithm and the baselines in §3.5.1, and then describe our main results in §3.5.2–3.5.5.

3.5.1. Phase I: Inferred Tips

Normal configuration. 183 participants[5] successfully completed the game in the first phase of the normal configuration. The top three tips inferred from each of the sources are reported in Table 8. For the algorithm tip, "Chef should never plate" is selected as it is expected to be the most effective at shortening completion time (2.43 ticks). For the baseline tip, "Chef should chop once" is chosen as it is the most frequently observed state-

[5] The average age of the participants is 34.6 years old, 57.38% are female, and 67.73% have at least a two-year degree.

action pair in the data. Finally, for the human tip, "Strategically leave some workers idle" received the most votes among previous participants (28.42%).

Table 8: Top three tips inferred from different sources for the normal configuration.

Normal	**Tip #1**	Tip #2	Tip #3
Algorithm	**Chef should never plate**	Server plates three times	Server should skip chopping once
Baseline	**Chef should chop once**	Server should plate three times	Sous-chef should plate twice
Human (% voted)	**Strategically leave some workers idle** (28.42%)	Server should never cook (21.31%)	Chef should never plate (13.11%)

It is worth noting that all of the top three human-generated tips are in line with the optimal policy. The first tip captures the key strategy that some virtual workers should be left idle rather than assigned to a time-consuming task. However, it is less specific than the other tips or tips from our algorithm. The second and third tips reflect the fact that the server takes the most time cooking and the chef takes the most time plating, which participants could learn from assigning these tasks to them during the game.

Disrupted configuration. 172 participants[6] successfully completed the first phase of the disrupted configuration. Table 9 reports the top three tips inferred from each of the sources. The best algorithm tip is "Server should cook twice" with the expected completion time reduction of 2.32 ticks. Interestingly, only the first and third tips are in line with the optimal policy. The second tip is slightly off as in the optimal policy Sous-chef and Server should each plate twice. For the baseline arm, all three tips are in line with the optimal policy and we select "Sous-chef should plate twice" as the baseline tip. Finally, for the human tip, we select the one with the most votes from the participants: "Server should cook once" or "Sous-chef should cook three times".

Unlike the top human tips in the normal configuration, the two tips with the most votes in the disrupted configuration are not part of the optimal policy. In the optimal policy, Sous-chef and Server should each cook twice. While the third human tip is in line with the optimal policy as no worker should be left idle in the disrupted scenario, it is much less

[6]They are 36.4 years old on average, 61.63% are female, and 77.91% received at least a two-year degree.

Table 9: Top three tips inferred from different sources for the disrupted configuration.

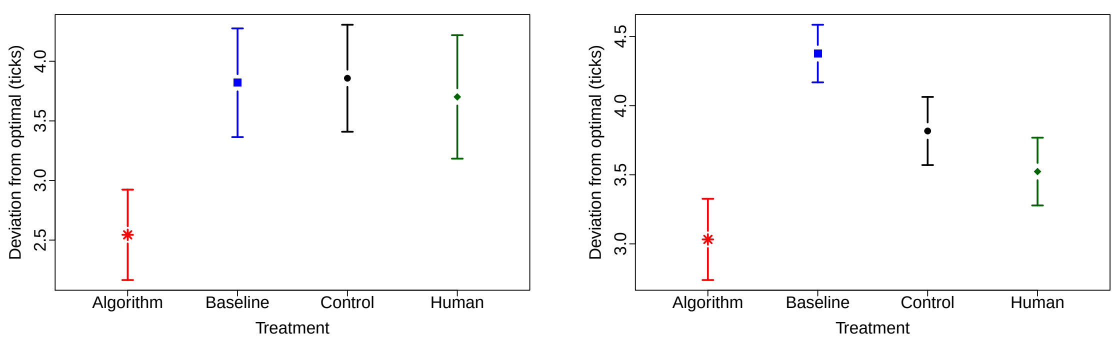
Disrupted	Tip #1	Tip #2	Tip #3
Algorithm	**Server should cook twice**	Sous-chef should plate once	Server should chop once
Baseline	**Sous-chef should plate twice**	Sous-chef should chop three times	Server should cook twice
Human (% voted)	**Server should cook once** (28.48%)	Server should never cook (23.84%)	Keep everyone busy (16.86%)

specific than the first two tips or those inferred from our algorithm. This highlights the increased difficulty for humans to identify the optimal strategy in the disrupted configuration compared to the normal configuration.

3.5.2. Phase II: Our Tips Substantially Improve Performance

Normal configuration. 1,317 participants[7] successfully completed the game. Participants shown our tip completed the final round in 22.54 steps on average (optimal is 20 steps), significantly outperforming those in other arms: 23.86 (control group, $t = -4.397$, $p < 0.0001$), 23.73 (human tip, $t = 3.628$, $p = 0.0002$), and 23.82 (naïve algorithm, $t = -4.232$, $p < 0.0001$); see Figure 20a.

Figure 20: Performance of participants across conditions in the last round of Phase II.



(a) Normal configuration (optimal: 20 ticks) (b) Disrupted configuration (optimal: 34 ticks)

Disrupted configuration. 1,011 participants[8] successfully completed the game. 244 were assigned to the control arm, 247 to the algorithm arm, 249 to the human arm, and 271 to the baseline arm. Participants shown our tip completed the final rounds in 37.05 steps,

[7]The mean age is 33.3 years old, 51.03% are female, and 67.73% completed a degree beyond high school.
[8]On average, participants are 34.9 years old, 60.14% are female, and 70.43% have at least a two-year degree.

again significantly outperforming those in other arms: 37.92 (control group, $t = -4.361$, $p < 0.0001$), 37.53 (human tip, $t = -2.52$, $p = 0.0061$), and 38.40 (naïve algorithm, $t = -7.348$, $p < 0.0001$); see Figure 20b.

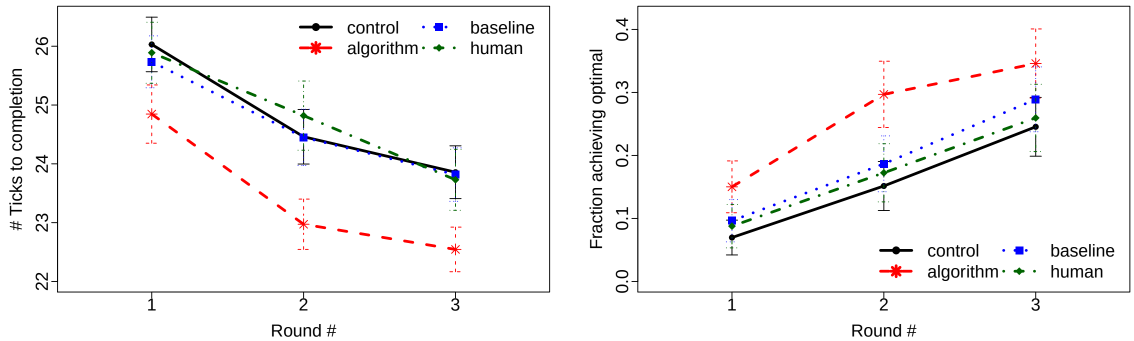
3.5.3. Learning Over Time: Our Tips Speed Up Learning

Next, we study how performance improves across rounds as participants learn better strategies. In particular, tips can be thought of as a substitute for learning, reducing the number of rounds needed for participants to achieve a certain performance level. Recall that participants in the normal configuration had three game rounds over which they could learn and improve, while those in the disrupted configuration had four rounds (not counting the initial two rounds with the fully-staffed scenario). In addition to the number of ticks to complete all orders, we also examine the fraction of participants that reach the optimal reward—i.e., completing the game in 20 ticks for the fully-staffed scenario or in 34 ticks for the understaffed scenario).

Normal configuration. Our tip speeds up learning by at least one round compared to any of the other arms. Figure 21a illustrates the performance measured by the number of ticks taken to complete each round for participants in each condition. First, we observe that participants in all arms improved over the three rounds. Although participants in each of the three arms (but not the control) received tips starting from the first round, only those receiving our tip performed significantly better the control in any of the three rounds. In fact, our tip speeds up learning by at least one round compared to any of the other arms—i.e., the performance of participants given our tip on the kth round was similar to the performance of participants in alternative arms on the $(k + 1)$th round. Figure 21b displays the fraction achieving this optimal performance for each arm over the rounds. As before, we observe that all participants improved over the three rounds, and that our tip was the only one to significantly improve performance in any round.

Disrupted configuration. The improved speed of learning was even more apparent under the disrupted configuration. Figure 22a shows the performance measured by the completion time in each of the under-staffed rounds for participants in each arm. The participants

Figure 21: Performance of participants in each condition across the rounds of Phase II (normal configuration).



(a) Ticks to completion

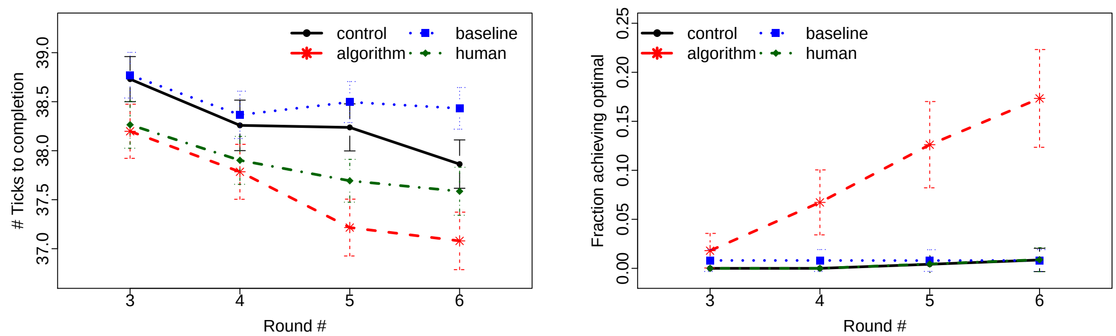
(b) Fraction achieving optimal

in the control arm took four rounds to achieve the same level of performance as the participants shown our tip on the first round. Thus, our optimal tip speeds up learning by four rounds. In addition, Figure 22b shows the fraction of participants that achieved optimal performance. Interestingly, only participants given our tip achieved optimal performance: by the last round of the game, 18.79% of participants in the algorithm arm played optimally, compared to 1.14% of the human tip arm, 0.99% of the baseline arm, and 0.51% of the control arm.

Intuitively, the baseline algorithm performs poorly since it blindly tries to mimic the optimal policy rather than focus on accounting for consequential decisions. For instance, in the disrupted configuration, it infers the tip "Sous-chef should plate twice", which actually reduces performance. The actions suggested by this tip occur at the end of the game, which is too late to significantly benefit overall performance. In contrast, our algorithm focuses on mimicking decisions made by the optimal policy that have long-term benefits.

Next, note that the human arm initially improves performance comparably to our arm; however, it levels off towards the end whereas our arm continues to improve. These results suggest that our tip, while simple and concise, encodes a complex underlying strategy that the participants come to understand when they combine it with their own experience playing the game. In contrast, the human-suggested tip encodes a more shallow strategy that quickly improves performance but does not lead to deeper insight over time.

Figure 22: Performance of participants in each condition in the last four rounds of Phase II (disrupted configuration).



(a) Ticks to completion

(b) Fraction achieving optimal

Lastly, we believe that one potential obstacle to uncovering the optimal strategy during the understaffed scenario is the participants' inability to "un-learn" and adapt their strategy to the disrupted environment. In the first two rounds with the fully-staffed scenario, participants have potentially learned each worker's skill level and developed a strategy to assign tasks based on such knowledge. A long task (e.g., cooking burger) is often assigned to the highly skilled worker (e.g., chef), while the least skilled worker (e.g., server) is reserved to carry on a short task (e.g., plating). Once the disruption took place, the majority of participants kept their original strategy—i.e., not assigning chopping or cooking tasks to the server. After four rounds of the understaffed scenario, a fraction of participants learned that leaving the server idle was suboptimal. The human-proposed tip ("Server should cook once") suggests that participants were able to adjust their strategy towards the optimal one after four rounds. However, this tip is not aggressive enough to achieve the optimal performance—as indicated by our tip ("Server should cook twice"). In optimal play, the server actually needs to perform a significantly larger share of the subtasks than the human tip suggests. Another evidence for this explanation is observed in the post-game survey of both phases of the normal configuration. Although participants in these studies did not experience a disruption, they were asked to imagine a hypothetical understaffed scenario and select the best tip that they expected to help improve performance in such disruption.

The tip that received the most votes is “Server shouldn’t cook”. Without the actual experience of managing the disruption, participants appeared to be biased towards their strategy learned in the normal scenario. Thus, we believe the success of our tip is due in part to how it helps human decision-makers overcome their resistance to exploring counterintuitive strategies.

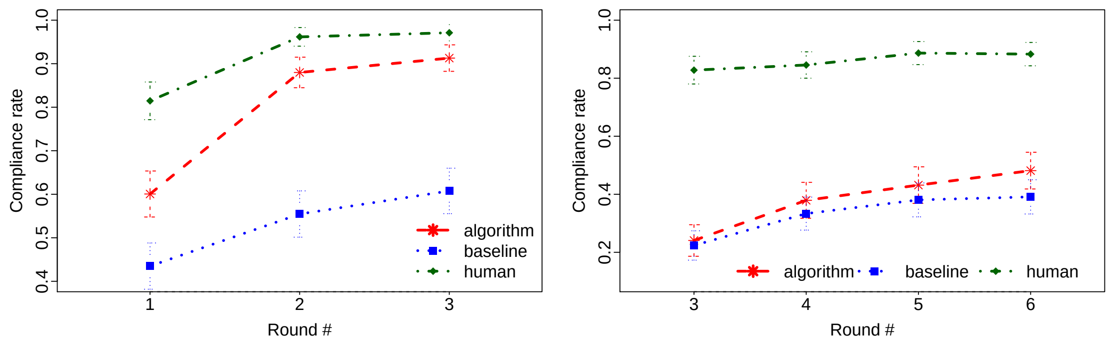
3.5.4. Complying to Tips: Human Users are More Compliant Over Time

The effectiveness of a tip critically depends on whether the participant follows it; to better understand this relationship, we study how well participants complied with tips across arms. Importantly, participants were not informed of the source of the tips, so variation in compliance is entirely due to the contents of the tips.

Normal configuration. We find that participants increasingly comply with tips across rounds in all arms, as can be seen in Figure 23a. However, in the final round, a significantly higher fraction participants complied with our tips and the human suggested tips compared to the baseline algorithm tips, suggesting that participants determined that the naïve algorithm did not suggest a useful tip. These results suggest that participants are not blindly following tips; instead, they only follow the tip if it suggests a strategy that makes sense to them. Furthermore, they show that our strategy is effective even though our tips are inferred under the assumption that the participant exactly follows the tip. Intuitively, we believe our approach remains effective since our objective of identifying a tip that maximizes long-term payoff is consistent with the idea that participants only follow the tip if it encodes an effective strategy; that is, they follow the strategy as long as they can understand it and it is effective.

Disrupted configuration. As before, Figure 23b shows that participants increasingly comply to their respective tips across rounds. The compliance of the human-suggested tip (“Server cooks once”) is significantly higher than the others, likely because it is the most intuitive. In contrast, our tip (“Server cooks twice”) is counterintuitive since the server is slow at cooking; nevertheless, it is more effective at improving performance when followed.

Figure 23: Compliance rate across the rounds of Phase II.



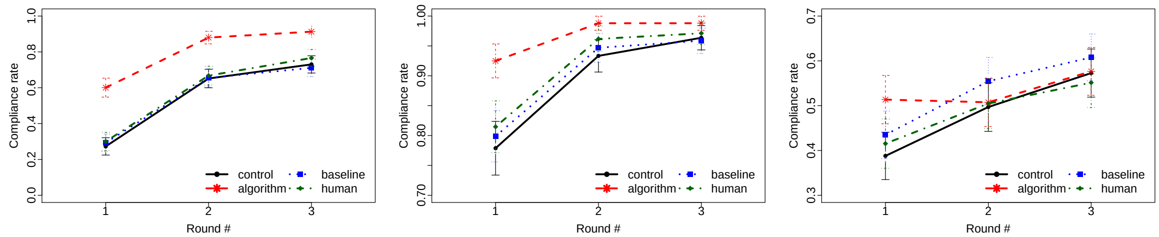
(a) Normal configuration (b) Disrupted configuration

3.5.5. Learning Beyond Tips: Our Tips Help Users Learn to Play Optimally

A key question is understanding how humans internalized and actualized the strategies encoded in the tips we inferred. We study this question in two ways. First, we examine *cross-compliance*, which is the compliance of the participant to alternative tips other than the one we showed them. Naïvely, there is no reason to expect participants to cross-comply with an alternative tip (assuming it does not overlap with the tip shown), beyond the cross-compliance of the control group to that tip. Thus, cross-compliance measures how showing one tip can enable participants to discover strategies beyond what is stated in that tip. Finally, we investigate whether the tips could help participants uncover the structure of the optimal policy beyond the simple rules they stated.

Cross-compliance. We find that our tips had high cross-compliance than other tips in both configurations. For the normal configuration, we find that participants across all arms learn not to assign plating to the chef (Figure 24a), strategically leave some virtual workers idle (Figure 24b), and let the chef chop only once (Figure 24c). These tips are all consistent with the optimal policy, suggesting that participants generally learn over time to improve their performance regardless of the treatment. Interestingly, participants in the algorithm arm have similar or higher cross-compliance compared to the other arms. This result suggests that our tip is the most effective as the information it conveys encompasses the information conveyed by the other tips.

Figure 24: Cross-compliance rate across the rounds of Phase II (normal configuration).



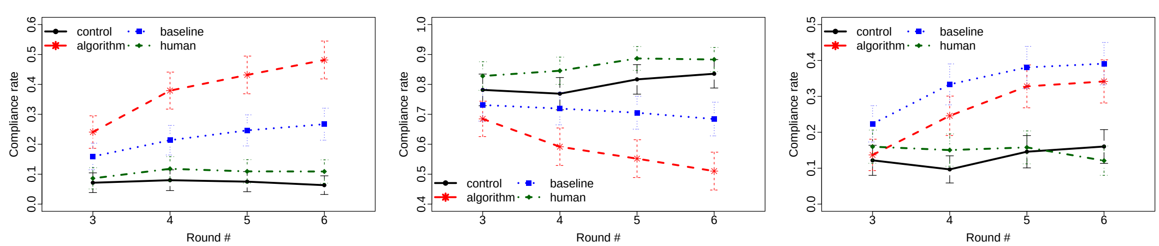
(a) Algorithm:
"Chef shouldn't plate"

(b) Human:
"Leave some idle"

(c) Baseline:
"Chef chops once"

For the disrupted configuration, the cross-compliance of the the human and control arms with our tip remained flat (Figure 25a). We observe a similar trend for the naïve algorithm tip (Figure 25c). These results suggest that participants do not naturally learn this strategy over time, most likely since it is counterintuitive. In particular, simple tips can greatly improve human performance by capturing counterintuitive strategies that take a great deal of experimentation to discover. Finally, Figure 25b shows the cross-compliance to the human tip. Interestingly, compliance of the algorithm arm to this strategy actually decreases over time; indeed, the tip suggested by humans is a suboptimal strategy, so complying with this tip leads to worse performance.

Figure 25: Cross-compliance rate across the rounds of Phase II (disrupted configuration).



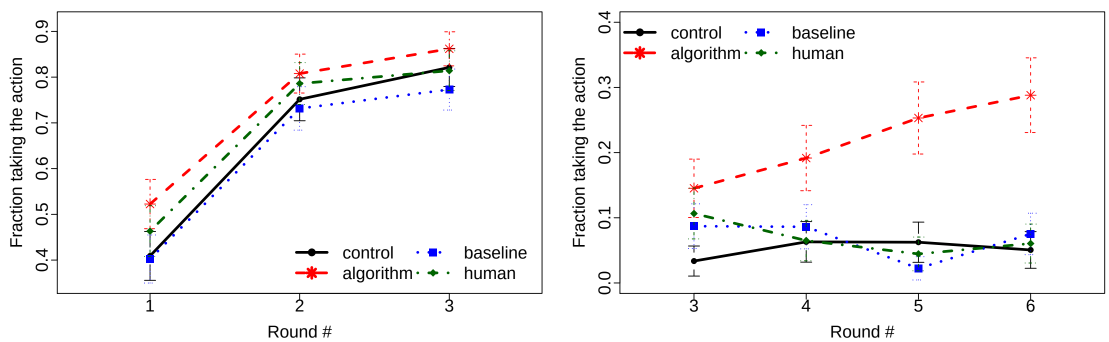
(a) Algorithm:
"Server cooks twice"

(b) Human:
"Server cooks once"

(c) Baseline:
"Sous-chef plates twice"

Uncovering the optimal policy. At a high level, the optimal policy for the fully-staffed scenario has the chef cook most of the dishes, has the server plate most of the dishes, and never assigns the chef to plate or the server to cook. We observe that participants generally recovered these optimal strategies as they played more rounds. For instance, the fraction of participants in each arm that never assigned cooking to the server in each round, as if they were following the tip "Server shouldn't cook", increases over time and within each round the fractions are not statistically different among the arms (see Figure 26a). This result suggests that people learned about this rule by themselves across all arms.

Figure 26: Fraction of participants taking optimal action beyond the tips they were shown across the rounds.



(a) Fully-staffed: "Server shouldn't cook"

(b) Under-staffed: "Server chops once"

Furthermore, for the under-staffed scenario, the optimal policy requires balanced assignment of cooking and plating tasks to both sous-chef and server while assigning most of the chopping tasks to sous-chef. A key rule beyond our tip is thus "Server chops once". We find that only participants in our arm cross-complied with this auxiliary tip (see Figure 26b), again demonstrating that our tip enables participants to discover strategies beyond what is stated in the tip.

3.6. Concluding Remarks

We have proposed a novel machine learning algorithm for automatically identifying interpretable tips designed to help improve human decision-making. Our behavioral study demonstrates that the tips inferred using our algorithm can successfully improve human performance on a challenging sequential decision-making task. In particular, our results for the normal configuration suggest that our tip can speed up learning by up to three rounds of in-game experience, demonstrating that our tip can significantly reduce the cost of learning. Furthermore, in the disrupted configuration, our results suggest that our tip enables the human participants to discover additional strategies beyond the tip. In other words, the benefit of tips comes not just from having the human follow the letter of the tip, but from how the human builds on the tip to discover additional insights.

An important ingredient in our framework is the incorporation of trace data to identify pieces of information that are most likely to help improve the performance of an average worker. Modern-day organizations have benefited from using customer data to inform new product strategies and provide personalized offerings to their customers, but the data on their own employees is underused. Our framework provides techniques to leverage the largely untapped potential of readily available trace data in pinpointing areas of performance improvement and identifying new practices. Even when the true optimal strategy is unknown, trace data of workers with high experience or good performance can be used to identify good strategies. In recent years, a growing number of organizations have adopted a gig economy employment model or allowed for remote work in response to worker preferences for flexibility and independence. To compensate for the lack of interactions among workers, firms can employ our algorithm to learn best practices from the highly performing workers and then provide tips to help individuals improve.

There are several important directions for future work. First, incorporating personalization to individual workers could greatly improve the performance of our tip inference algorithm. Our optimal tips were chosen based on the expected improvement among the bottom quartile of performers in the game. Ideally, we would instead infer tips personalized for different skill levels and individual worker characteristics. Furthermore, in our approach, we only inferred tips at one point in time. In practice, our approach could also be performed every time additional data is collected. Then, an important question is understanding the long-term benefits of our approach and understanding how it affects learning behavior over a longer period of time. Another promising direction is extending our algorithm to a collaborative setting. We have only studied how individual workers learn to improve performance, but a similar approach may help teams improve their collaboration and optimize information sharing. Finally, future work is needed to study how to better convey machine-generated tips to improve compliance. Recent work has documented human aversion to advice made by algorithms (Eastwood et al. 2012, Dietvorst et al. 2015) and shown certain conditions that alleviate such aversion (Dietvorst et al. 2018, Logg et al. 2019). In our study, a fraction of participants chose to forgo our tip and continue using their own strategy. Finding ways to build trust and encourage compliance is an important ingredient for ensuring our tips help people improve.

BIBLIOGRAPHY

P. Abbeel and A. Y. Ng. Apprenticeship learning via inverse reinforcement learning. In *Proceedings of the twenty-first international conference on Machine learning*, page 1, 2004.

S. Agarwal, M. Diao, J. Pan, and T. F. Sing. Are singaporean cabdrivers target earners? Working paper, 2015.

Z. Akşin, S. Deo, J. O. Jónasson, and K. Ramdas. Learning from many: Partner exposure and team familiarity in fluid teams. *Management Science*, 2020.

L. Argote. *Organizational learning: Creating, retaining and transferring knowledge*. Springer Science & Business Media, 2012.

M. Arvan, B. Fahimnia, M. Reisi, and E. Siemsen. Integrating human judgement into quantitative forecasting methods: A review. *Omega*, 86:237–252, 2019.

A. Asadpour, I. Lobel, and G. van Ryzin. Minimum earnings regulation and the stability of marketplaces. *Available at SSRN 3502607*, 2019.

O. Bastani, Y. Pu, and A. Solar-Lezama. Verifiable reinforcement learning via policy extraction. In *Advances in neural information processing systems*, pages 2494–2504, 2018.

H. Bavafa and J. O. Jónasson. Recovering from critical incidents: Evidence from paramedic performance. *Manufacturing & Service Operations Management*, 2020a.

H. Bavafa and J. O. Jónasson. The variance learning curve. *Available at SSRN*, 2020b.

D. G. Beckers, D. van der Linden, P. G. Smulders, M. A. Kompier, T. W. Taris, and S. A. Geurts. Voluntary or involuntary? control over overtime and rewards for overtime in relation to fatigue and work satisfaction. *Work & Stress*, 22(1):33–50, 2008.

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

S. Benjaafar, J.-Y. Ding, G. Kong, and T. Taylor. Labor welfare in on-demand service platforms. *Available at SSRN 3102736*, 2019.

S. Benjaafar, S. Xiao, and X. Yang. Do workers and customers benefit from competition between on-demand service platforms? *Available at SSRN*, 2020.

F. Bourguignon, M. Fournier, and M. Gurgand. Selection bias corrections based on the multinomial logit model: Monte carlo comparisons. *Journal of Economic Surveys*, 21(1):174–205, 2007.

T. Brachet, G. David, and A. M. Drechsler. The effect of shift structure on performance. *American Economic Journal: Applied Economics*, 4(2):219–46, 2012.

H. Brattland, J. R. Høiseth, O. Burkeland, T. S. Inderhaug, P. E. Binder, and V. C. Iversen. Learning from clients: A qualitative investigation of psychotherapists' reactions to negative verbal feedback. *Psychotherapy Research*, 28(4):545–559, 2018.

R. L. Bray, J. C. Serpa, and A. Colak. Supply chain proximity and product quality. *Management science*, 65(9):4079–4099, 2019.

L. Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

K. A. Bryan and J. S. Gans. A theory of multihoming in rideshare competition. *Journal of Economics & Management Strategy*, 28(1):89–96, 2019.

G. P. Cachon, K. M. Daniels, and R. Lobel. The role of surge pricing on a service platform with self-scheduling capacity. *Manufacturing & Service Operations Management*, 19(3):368–384, 2017.

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

W. J. Carrington. The alaskan labor market during the pipeline era. *Journal of Political Economy*, 104(1): 186–218, 1996.

T. Y. Chan, J. Li, and L. Pierce. Learning from peers: Knowledge transfer and sales force productivity growth. *Marketing Science*, 33(4):463–484, 2014.

M. K. Chen and M. Sheldon. Dynamic pricing in a labor market: Surge pricing and flexible work on the uber platform. In *EC*, page 455, 2016.

M. K. Chen, P. E. Rossi, J. A. Chevalier, and E. Oehlsen. The value of flexible work: Evidence from uber drivers. *Journal of Political Economy*, 127(6):2735–2794, 2019.

Y.-J. Chen, T. Dai, C. G. Korpeoglu, E. Körpeoğlu, O. Sahin, C. S. Tang, and S. Xiao. Om forum—innovative online platforms: Research opportunities. *Manufacturing & Service Operations Management*, 22(3):430–445, 2020.

M. Chui, J. Manyika, and J. Bughin. The social economy: Unlocking value and productivity through social technologies. Technical report, McKinsey Global Institute, 2012.

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

B. Corgnet, S. Gaechter, and R. Hernán-González. Working too much for too little: stochastic rewards cause work addiction. *Available at SSRN 3540225*, 2020.

J. G. Cragg. Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica (pre-1986)*, 39(5):829, 1971.

V. P. Crawford and J. Meng. New york city cab drivers' labor supply revisited: Reference-dependent preferences with rational-expectations targets for hours and income. *American Economic Review*, 101(5): 1912–32, 2011.

M. Csikszentmihalyi and I. S. Csikszentmihalyi. *Optimal experience: Psychological studies of flow in consciousness*. Cambridge university press, 1992.

M. Csikszentmihalyi and J. LeFevre. Optimal experience in work and leisure. *Journal of personality and social psychology*, 56(5):815, 1989.

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

R. M. Dawes, D. Faust, and P. E. Meehl. Clinical versus actuarial judgment. *Science*, 243(4899):1668–1674, 1989.

S. E. DeVoe and J. Pfeffer. When time is money: The effect of hourly payment on the evaluation of time. *Organizational Behavior and Human Decision Processes*, 104(1):1–13, 2007.

S. E. DeVoe, B. Y. Lee, and J. Pfeffer. Hourly versus salaried payment and decisions about trading time and money over time. *ILR Review*, 63(4):627–640, 2010.

G. Dhaene and K. Jochmans. Split-panel jackknife estimation of fixed-effect models. *The Review of Economic Studies*, 82(3):991–1030, 2015.

B. J. Dietvorst, J. P. Simmons, and C. Massey. Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1):114, 2015.

B. J. Dietvorst, J. P. Simmons, and C. Massey. Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. *Management Science*, 64(3):1155–1170, 2018.

J. Dong and R. Ibrahim. Managing supply in the on-demand economy: Flexible workers, full-time employees, or both? *Operations Research*, 2020.

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

B. Dorn and M. Guzdial. Learning on the job: characterizing the programming knowledge and learning strategies of web designers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 703–712, 2010.

J. Eastwood, B. Snook, and K. Luther. What people want from their professionals: Attitudes toward decision-making strategies. *Journal of Behavioral Decision Making*, 25(5):458–468, 2012.

J. A. Evans, G. Kunda, and S. R. Barley. Beach time, bridge time, and billable hours: The temporal structure of technical contracting. *Administrative Science Quarterly*, 49(1):1–38, 2004.

H. S. Farber. Is tomorrow another day? the labor supply of new york city cabdrivers. *Journal of political*

Economy, 113(1):46–82, 2005.

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

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

V. Farewell, D. Long, B. Tom, S. Yiu, and L. Su. Two-part and related regression models for longitudinal data. *Annual review of statistics and its application*, 4:283–315, 2017.

E. G. Fitzsimmons. Uber hit with cap as new york city takes lead in crackdown, Aug 2018. URL `https://www.nytimes.com/2018/08/08/nyregion/uber-vote-city-council-cap.html`.

M. Gleicher. A framework for considering comprehensibility in modeling. *Big data*, 4(2):75–88, 2016.

C. Gourieroux, A. Monfort, and E. Renault. Indirect inference. *Journal of applied econometrics*, 8(S1): S85–S118, 1993.

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

I. Gurvich, K. J. O'Leary, L. Wang, and J. A. Van Mieghem. Collaboration, interruptions, and changeover times: Workflow model and empirical study of hospitalist charting. *Manufacturing & Service Operations Management*, 2019.

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

J. J. Heckman. Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, pages 153–161, 1979.

K. Herkenhoff, J. Lise, G. Menzio, and G. Phillips. Knowledge diffusion in the workplace. Technical report, July 2018. Working Paper, University of Minnesota, 2018.

M. Hu and Y. Zhou. Price, wage and fixed commission in on-demand matching. *Available at SSRN 2949513*, 2019.

R. S. Huckman and G. P. Pisano. The firm specificity of individual performance: Evidence from cardiac surgery. *Management Science*, 52(4):473–488, 2006.

R. G. Hughes et al. Patient safety and quality. *an evidence-based handbook for nurses*, page 2008, 2008.

M. R. Ibanez, J. R. Clark, R. S. Huckman, and B. R. Staats. Discretionary task ordering: Queue management in radiological services. *Management Science*, 64(9):4389–4407, 2017.

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

R. Ibrahim, S.-H. Kim, and J. Tong. Eliciting human judgment for prediction algorithms. *Available at SSRN*, 2020.

G. Imbens and J. Wooldridge. Control function and related methods. What's new in Econometrics, National Bureau of Economic Research, 2007.

E. Intelligence. Freelancing in america: 2019, 2019.

G. Jarosch, E. Oberfield, and E. Rossi-Hansberg. Learning from coworkers. Technical report, National Bureau of Economic Research, 2019.

A. Kabra, B. Elena, and G. Karan. The efficacy of incentives in scaling marketplaces. Working paper, 2017.

T. Kaji, E. Manresa, and G. Pouliot. An adversarial approach to structural estimation. *arXiv preprint 2007.06169*, 2020.

N. Karacaoglu, A. Moreno, and C. Ozkan. Strategically giving service: The effect of real-time information on service efficiency. *Available at SSRN 3260035*, 2018.

D. S. Kc and B. R. Staats. Accumulating a portfolio of experience: The effect of focal and related experience on surgeon performance. *Manufacturing & Service Operations Management*, 14(4):618–633, 2012.

S.-H. Kim, H. Song, and M. Valentine. Staffing temporary teams: Understanding the effects of team familiarity and partner variety. *Available at SSRN 3176306*, 2020.

J. Kleinberg, J. Ludwig, S. Mullainathan, and Z. Obermeyer. Prediction policy problems. *American Economic Review*, 105(5):491–95, 2015.

D. Kruzman. Some uber drivers work dangerously long shifts, Jul 2017. URL `https://www.usatoday.com/story/money/cars/2017/07/10/some-uber-drivers-work-dangerously-long-shifts/103090682/`.

F. Leclerc, B. H. Schmitt, and L. Dube. Waiting time and decision making: Is time like money? *Journal of Consumer Research*, 22(1):110–119, 1995.

B. Letham, C. Rudin, T. H. McCormick, D. Madigan, et al. Interpretable classifiers using rules and bayesian analysis: Building a better stroke prediction model. *The Annals of Applied Statistics*, 9(3):1350–1371, 2015.

J. Levinsohn and A. Petrin. Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies*, 70(2):317–341, 2003.

Q. Liu, O. Loginova, and X. H. Wang. The impact of multi-homing in a ride-hailing market. *Available at SSRN 2968504*, 2017.

J. M. Logg, J. A. Minson, and D. A. Moore. Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, 151:90–103, 2019.

D. Madden. Sample selection versus two-part models revisited: The case of female smoking and drinking. *Journal of health economics*, 27(2):300–307, 2008.

J. Manyika, S. Lund, K. Robinson, J. Valentino, and R. Dobbs. A labor market that works: Connecting talent with opportunity in the digital age. *June. McKinsey Global Institute. http://www. mckinsey. com/˜/media/McKinsey/dotcom/Insights/Employment% 20and% 20growth/Connecting*, 20, 2015.

A. Mao, W. Mason, S. Suri, and D. J. Watts. An experimental study of team size and performance on a complex task. *PloS one*, 11(4):e0153048, 2016.

A. Marshall. Uber changes its rules, and drivers adjust their strategies, Feb 2020. URL `https://www.wired.com/story/uber-changes-rules-drivers-adjust-strategies/`.

C. H. Mason and W. D. Perreault Jr. Collinearity, power, and interpretation of multiple regression analysis. *Journal of marketing research*, 28(3):268–280, 1991.

C. K. Morewedge, H. Yoon, I. Scopelliti, C. W. Symborski, J. H. Korris, and K. S. Kassam. Debiasing decisions: Improved decision making with a single training intervention. *Policy Insights from the Behavioral and Brain Sciences*, 2(1):129–140, 2015.

W. J. Murdoch, C. Singh, K. Kumbier, R. Abbasi-Asl, and B. Yu. Definitions, methods, and applications in interpretable machine learning. *Proceedings of the National Academy of Sciences*, 116(44):22071–22080, 2019.

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

I. Nonaka and H. Takeuchi. *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. Oxford university press, 1995.

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

E. M. Okada and S. J. Hoch. Spending time versus spending money. *Journal of consumer research*, 31(2): 313–323, 2004.

J. A. Parrott and M. Reich. An earnings standard for new york city's app-based drivers. *New York: The New School: Center for New York City Affairs*, 2018.

J. Pfeffer, R. I. Sutton, et al. *The knowing-doing gap: How smart companies turn knowledge into action.* Harvard business press, 2000.

P. Puhani. The heckman correction for sample selection and its critique. *Journal of economic surveys*, 14 (1):53–68, 2000.

K. Ramdas, K. Saleh, S. Stern, and H. Liu. Variety and experience: Learning and forgetting in the use of surgical devices. *Management Science*, 64(6):2590–2608, 2017.

N. Rosaia. Competing platforms and transport equilibrium: Evidence from new york city. Technical report, mimeo, Harvard University, 2020.

C. Rudin. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5):206–215, 2019.

A.-L. Sellier, I. Scopelliti, and C. K. Morewedge. Debiasing training improves decision making in the field. *Psychological science*, 30(9):1371–1379, 2019.

M. Sheldon. Income targeting and the ridesharing market. Working paper, 2016.

S. A. Sisson, Y. Fan, and M. M. Tanaka. Sequential monte carlo without likelihoods. *Proceedings of the National Academy of Sciences*, 104(6):1760–1765, 2007.

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

H. Song, A. L. Tucker, K. L. Murrell, and D. R. Vinson. Closing the productivity gap: Improving worker productivity through public relative performance feedback and validation of best practices. *Management Science*, 64(6):2628–2649, 2017.

S. J. Spear. Fixing health care from the inside, today. *Harvard business review*, 83(9):78, 2005.

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

D. N. Sull and K. M. Eisenhardt. *Simple rules: How to thrive in a complex world*. Houghton Mifflin Harcourt, 2015.

R. S. Sutton, D. A. McAllester, S. P. Singh, and Y. Mansour. Policy gradient methods for reinforcement learning with function approximation. In *Advances in neural information processing systems*, pages 1057–1063, 2000.

G. Szulanski. Exploring internal stickiness: Impediments to the transfer of best practice within the firm. *Strategic management journal*, 17(S2):27–43, 1996.

T. F. Tan and S. Netessine. When you work with a superman, will you also fly? an empirical study of the impact of coworkers on performance. *Management Science*, 65(8):3495–3517, 2019.

T. A. Taylor. On-demand service platforms. *Manufacturing & Service Operations Management*, 20(4): 704–720, 2018.

N. Thakral and L. T. Tô. Daily labor supply and adaptive reference points. American Economic Review, 2019.

A. L. Tucker, A. C. Edmondson, and S. Spear. When problem solving prevents organizational learning. *Journal of Organizational Change Management*, 2002.

A. L. Tucker, I. M. Nembhard, and A. C. Edmondson. Implementing new practices: An empirical study of organizational learning in hospital intensive care units. *Management science*, 53(6):894–907, 2007.

M. Watanabe and K. Yamauchi. Psychosocial factors of overtime work in relation to work-nonwork balance: A multilevel structural equation modeling analysis of nurses working in hospitals. *International journal of behavioral medicine*, 23(4):492–500, 2016.

C. J. Watkins and P. Dayan. Q-learning. *Machine learning*, 8(3-4):279–292, 1992.

S. Weisband. Maintaining awareness in distributed team collaboration: Implications for leadership and performance. *Distributed work*, pages 311–333, 2002.

R. Wilson. Seattle becomes first city to cap uber, lyft vehicles, Apr 2019. URL `https://www.washingtonpost.com/blogs/govbeat/wp/2014/03/18/seattle-becomes-first-city-to-cap-uber-lyft-vehicles/`.

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